

Studenters opplevelse og evaluering av undervisning og læringsmiljø: Presentasjon av Course Experience Questionnaire (CEQ) – og validering av tre norske versjoner, Erfaringer med studiet (EMS)

Roar C. Pettersen

**Høgskolen i Østfold
Rapport 2007:4**

Online-versjon (pdf)

Utgivelsessted: Halden

Det må ikke kopieres fra rapporten i strid med åndsverkloven og fotografiloven eller i strid med avtaler om kopiering inngått med KOPINOR, interesseorgan for rettighetshavere til åndsverk.

Høgskolen i Østfold har en godkjenningsordning for publikasjoner som skal gis ut i Høgskolens Rapport- og Arbeidsrapportserier.

Rapporten kan bestilles ved
henvendelse til Høgskolen i Østfold.
(E-post: postmottak@hiof.no)

Høgskolen i Østfold. Rapport 2007:4
© Forfatteren/Høgskolen i Østfold
ISBN: 978-82-7825-206-2
ISSN: 1503-2612

Forord

Bidrag fra flere personer har vært til hjelp for å ferdigstille denne rapporten; noen helt avgjørende: En takk til de nærmere tre tusen studentene som har deltatt i evaluering sine kurs, studieprogrammer og undervisnings- og læringsmiljøer i perioden 2000–2006, og som på denne måten har ”krevd sin rett – og gjort sin plikt” i kvalitetsreformens forstand.

Kvalitetsreformens anvisninger om at studentevaluering skal anvendes i lagt større utstrekning en tidligere, og etterlysningen av ”skikkelige verktøy” til dette formålet, er et av motivene for å foreta en bred og grundig validering av norske versjoner av evalueringsverktøyet *Course Experience Questionnaire*. I denne prosessen har kollegaer ved Høgskolen i Østfold, og ved andre høyskoler, deltatt på ulike måter: Både i administrasjon av spørreskjemaer og i fruktbare diskusjoner om studentevaluering generelt, og drøfting av EMS-resultater spesielt. En takk til dere.

En spesiell takk til professor Per Lauvås for nyttig feedback på tidlige manusutkast, og for samarbeidet med å introdusere og teste ut EMS som avdelingsovergripende evalueringsstrategi ved høyskolen.

Sist, men ikke minst – en takk til biblioteket ved Avd. for Helse- og sosialfag, HiØ, for at dere alltid skaffer til veie bøker, publikasjoner og artikler fra nær og fjern, raskt og effektivt.

Kråkerøy, den 26. februar 2007

Roar C. Pettersen

Innholdsfortegnelse

Innledning	7
DEL 1 Presentasjon av <i>Course Experience Questionnaire</i>	11
Utvikling av Course Experience Questionnaire (CEQ)	13
Lancasterstudiene	13
Et utvidet perspektiv på studenters læring	14
Fra ”course perception” til ”course experience”	18
Generelle ferdigheter og kompetanseinnretninger	22
Ulike versjoner av CEQ	23
EMS, norske varianter av CEQ	24
 DEL 2 Studentevaluering og CEQ/EMS i praksis Prinsipper, strategier og eksempler	 25
Studentevaluering – intensjoner og strategier	27
Formative og summative intensjoner	28
Studentevaluering som indikasjon på undervisningskvalitet	29
CEQ som utviklingsorientert evalueringsverktøy	31
Presentasjon og analyse av CEQ-data	31
Faktorskårer	32
Prosentvis enighet – frekvensfordeling	36
Presentasjon og analyse på program-/klassenivå – eksempler	37
 DEL 3 Validering av <i>Erfaringer med studiet</i> (EMS) Tre studier med norske versjoner av CEQ	 41
Validering av <i>Erfaringer med studiet</i>	43
Utvalgte valideringsstudier med CEQ – en kort oversikt	43
Validering av CEQ36	43
Validering av CEQ23/24	44

Kort om faktoranalyse og begrepsvaliditet	45
Studie A <i>Erfaringer med studiet</i> – versjon EMS30	49
Begrepsvaliditet	49
Reliabilitet – skalaenes indre konsistens	53
Studie B <i>Erfaringer med studiet</i> – versjon EMS30–GF	55
Begrepsvaliditet	56
Reliabilitet – indre konsistens	57
En merknad om <i>Generelle ferdigheter</i> og CEQ23.....	58
Studie C <i>Erfaringer med studiet</i> – versjon EMS26	61
Begrepsvaliditet og reliabilitet	62
EMS – en generell indikator for opplevd undervisningskvalitet?	63
EMS – kriterievaliditet	67
EMS som uttrykk for studentenes tilfredshet med studiet.....	67
EMS og læringstilnærminger/studiestrategier.....	68
Generelle ferdigheter som eksternt kriterium	73
Valideringsstudiene A, B og C oppsummert – konklusjon	75
Oppsummering – kort sluttord	79
Referanser	83
Vedlegg	89
Vedlegg 1 EMS30	90
Vedlegg 2 EMS30–GF	91
Vedlegg 3 EMS26	92
Vedlegg 4 Tabell: Skalalengde, inter–item korrelasjon og Cronbachs alfa	93

Tabelloversikt

Tabell 1	Dybdetilnærming og overflatetilnærming (oversikt)	16
Tabell 2	CEQ-/EMS-faktorskårer.....	33
Tabell 3	Rekodete faktorskårer	36
Tabell 4	Faktoranalyse, mønstermatrise for EMS30, n=1611.....	50
Tabell 5	Mønstermatrise for 17 utsagn fra EMS30 videreført i CEQ23	53
Tabell 6	Cronbachs alfa for skalaene i EMS30	53
Tabell 7	Mønstermatrise for EMS30–GF, n= 317	56
Tabell 8	Cronbachs alfa for EMS30–GF	57
Tabell 9	Samlet utvalg i Studie C, n=876	61
Tabell 10	Faktoranalyse, mønstermatrise for EMS26, n=876	62
Tabell 11	Cronbachs alfa for EMS26	63
Tabell 12	Annenordens faktoranalyse basert på faktorskårer i EMS	64
Tabell 13	Korrelasjon mellom skalaene i EMS/CEQ og tilfredshet	68
Tabell 14	Korrelasjon: Dybdetilnærming – EMS-/CEQ-skalaer	70
Tabell 15	Korrelasjon: Overflatetilnærming – EMS-/CEQ-skalaer	71
Tabell 16	Korrelasjon: Strategisk tilnærming – EMS-/CEQ-skalaer	72
Tabell 17	Korrelasjon: Generelle ferdigheter – øvrige EMS-/CEQ-skalaer	74

Innledning

The Course Experience Questionnaire (CEQ) har i løpet av 15–20 år blitt et internasjonalt kjent og anerkjent studentevalueringsinstrument for kartlegging av opplevd undervisningskvalitet og studenters vurdering av læringsmiljøer.

Rapportens første del presenterer CEQ og den teoretiske referanserammen spørreskjemaet er utviklet innenfor; en forskningstradisjon gjerne omtalt som *Student approaches to learning*, eller i kortform *SAL-tradisjonen*. I første del beskrives også ulike versjoner av CEQ og hovedstrukturen i de norske, oversatte versjonene, betegnet *Erfaringer med studiet*, forkortet EMS.

Del 2 har studentevaluering av undervisning som tema og drøfter så vel prinsipielle som praktiske forhold knyttet til bruken av *Erfaringer med studiet* som strategi i studentevaluering av undervisnings- og læringsmiljøer.

Innenfor høyere utdanning i Norge er studentevaluering av undervisning viet betydelig oppmerksomhet, særlig det siste ti–femten årene. Spørsmålet om studentevaluering ble reist i utredningen *Med viten og vilje* (NOU 1988:28), fulgt opp i St.meld. nr. 40 (1990–91) om høyere utdanning og av en arbeidsgruppe nedsatt av departementet i 1993, som la fram innstillingen *Studentevaluering av undervisning* (jf. Handal 1996). Senere påla departementet landets universiteter og høyskoler om å iverksette studentevaluering i løpet av studieåret 1994–95 og innførte i den forbindelse en evalueringspris på 500.000 kr, både som incitament og for å få fram evalueringsopplegg som kunne tjene som gode eksempler.

Kvalitetsreformen, innført fra studieåret 2003–04,¹ pålegger utdanningsinstitusjoner å utarbeide kvalitetssystemer hvor studentevalueringsrutiner inngår som element. Det nyopprettede NOKUT (Nasjonalt organ

¹ Med utgangspunkt i St.meld. nr.27 (2000-2001): *Gjør din plikt – Krev din rett. Kvalitetsreform av høyere utdanning*. Tilråding fra Kirke-, utdannings- og forskningsdepartementet av 9. mars 2001.

for kvalitet i utdanningen) har fått oppgaven med å evaluere og godkjenne institusjonenes kvalitetssystemer. Med andre ord er ikke studentevaluering av undervisning lenger en øvelse bare for spesielt interesserte lærer og pedagoger.

Prinsipielle sider ved studentevaluering av undervisning skisseres i del 2; her tematiseres ulike intensjoner som ligger til grunn for studentevaluering, og ulike strategier og metoder for systematisk kartlegging av studenters opplevelse og vurdering av undervisning og læringsmiljø. Et hovedpoeng er at CEQ/EMS på den ene siden kan inngå som verktøy i studentevaluering som primært sikter på *kvalitetssikring*, som element i institusjonelle kvalitetssystemer. På den andre siden kan spørreskjemaet med fordel brukes i evaluering hvor *kvalitetsutvikling* er primær intensjon; for eksempel av lærere og lærergrupper i forhold til egen undervisning og egne kurs. En presisering av skillet mellom *summativ* og *formativ* evaluering nyanserer og utdyper CEQs rolle i forhold til disse to anvendelsesområdene, særlig gjelder det CEQ/EMS anvendt som *komplementær strategi* i evalueringsarbeidet. I den sammenheng beskrives praktiske sider ved CEQ/EMS mer detaljert, både i forhold til administrasjon av spørreskjemaet og de vanligste måtene for presentasjon, analyse og tolkning av evalueringsresultater. Videre presenteres eksempler som presiserer ideen med CEQ/EMS som komplementær strategi, der kvalitetsutviklede strategier kobles til metoder med en mer utpreget summativ, kvalitetssikrende intensjon.

I rapportens tredje del presenteres resultater fra valideringsstudier med tre versjoner av EMS, hhv. studie A, B og C. Studiene utgjør rapportens empiriske hoveddel hvor prosjektet er å kryssvalidere CEQ: Å undersøke om norske versjoner (EMS) dokumenterer psykometriske egenskaper tilsvarende resultater vi finner i angloamerikanske og australske utdanningskontekster. En rekke prosjekter og forskningsarbeider viser CEQ som et instrument med tilfredsstillende validitet og reliabilitet, på tvers av fag- og undervisningsmiljøer og utdanningskulturer. I utgangspunktet kvalifiserer CEQ til kategorien ”skikkelige verktøy”, som kvalitetsreformen etterlyser (St.meld. nr. 27 (2000–2001):11).

Imidlertid bør ethvert spørreskjema og inventorie prinsipielt sett valideres fra grunnen før det anvendes i nye sammenhenger. Særlig gjelder det instrumenter som oversettes og anvendes i nye språklige kontekster. Det handler om å avklare spørsmål som gjelder de aktuelle skjemaene 'portability': "... that is, their appropriateness as research tools in different national systems of higher education" (Richardson 2004:353). For kontekstsensitive instrumenter, som CEQ, er det særlig viktig at de valideres i flere utdanningskontekster, også nasjonalt.

De tre studiene med *Erfaringer med studiet* sikter mot en bred kryss-validering av CEQ. Studiene bygger på data fra perioden 2000–2006 og omfatter komplette datasett fra totalt 2.804 studenter. CEQ-versjonene i studie B og C representerer i tillegg versjoner med en skalastruktur som ikke er eksplisitt validert tidligere. Resultater fra de tre studiene diskuteres i forhold valideringsstudier med CEQ foretatt i andre utdanningskontekster, i hovedsak ved australske og britiske universiteter og høyskoler.

DEL 1

Presentasjon av *Course Experience Questionnaire*

Utvikling av Course Experience Questionnaire (CEQ)

Lancasterstudiene

The Course Experience Questionnaire er en videreutvikling av spørreskjemaet "Course Perceptions Questionnaire" (CPQ) som ble utarbeidet i forbindelse med de såkalte Lancasterstudiene, to femårige forskningsprosjekter ved Universitetet i Lancaster gjennomført i perioden 1968–1981 (Brown & Atkins 1988). Det første prosjektet søkte å kartlegge typiske individuelle karakteristika og kjennetegn *ved studenter* som kan assosieres med og forklare graden av akademisk "suksess og fiasko". Prosjektets hovedfokus var på hvordan ulike faktorer som tidligere skoleresultater, evner, personlighetstrekk, motivasjon og studie- og arbeidsmåter kan forklare forskjeller i studenters læringsresultater og akademiske prestasjoner. Prosjektet er dokumentert i *Degrees of Excellence: The Academic Achievement Game* (Entwistle & Wilson 1977).

Perspektivet i dette prosjektet føyer seg inn i en pedagogisk forskningstradisjon der man i hovedsak søker å *forklare* og *predikere* akademiske prestasjoner basert på studenters og studentgruppers *individuelle forskjeller* og karakteristika (studentforutsetninger). Det vi si at man, i en viss forstand, undersøker forhold som nærmest er uavhengige av særtrekk og sentrale variasjoner i det undervisnings- og læringsmiljøet studentene deltar og arbeider i til daglig.

Typiske forskjeller i studentenes individuelle forutsetninger, som delvis forklarer graden av akademisk suksess, ble identifisert i den første Lancasterstudien. Imidlertid avdekket man et sentralt poeng i intervjuer med studentene som skulle danne hovedmoment i det andre Lancasterprosjektet; intervjumaterialet viste klart at "... the differing groups not only tackled their work in contrasting ways, but *had very different perceptions of the university as an academic and social environment*" (Entwistle 1987:15; min kursivering).

Den *andre* Lancasterstudien søker å beskrive og forklare mer inngående hvordan "different perceptions" – forstått som studenters kvalitativt forskjellige erfaringer med og opplevelser av det aktuelle lærings- og

undervisningsmiljøet – påvirker deres valg og utforming av kvalitativt forskjellige studiestrategier og læringsmønstre. Kort og godt, prosjektets fokus er på: "... the marked effects of the academic context on the quality in student learning" (ibid:20).

Et utvidet perspektiv på studenters læring

Perspektiv og forklaringsmodell endres i det andre Lancasterprosjektet; forståelsen av studenters læring utvides. Studenters *persepsjon* og *opplevelse av studiekontekst og læringsmiljø* etableres som en *kritisk* mellomliggende (medierende) variabel for å forklare forskjeller i studenters læringsresultater og akademiske prestasjoner. Man forlot en mer lineær og endimensjonal forklaringsmodell som skissert i punkt (1) under, og erstattet den med en modell hvor forholdet mellom *læring* og *undervisning* framstår som et mer komplekst interaksjonsforhold, som en *tre-veis relasjon* og prosess; skissert i punkt (2).

Under punkt (1) forstås studenters repertoar av studie- og læringsstrategier, arbeidsmetoder og læringsmønstre *hovedsakelig* som person- og individforankrede ("innebygde") egenskaper og kompetanser. Her kan kvalitative forskjeller observeres: noen studenter anvender mer effektive studie- og læringsstrategier enn andre. Det er disse kvalitative forskjellene som bidrar til å forklare variasjonen i læringsresultater og graden av akademisk suksess.

(1) Studentforutsetninger → studie- og læringsstrategier → læringsresultater

(2) Studentforutsetninger ← → oppfatninger/erfaringer (perceptions) av den konkrete, aktuelle lærings- og studiekonteksten



- forskjeller i studie- og læringsorientering
- bruk av ulike studie- og læringsstrategier



læringsresultater

Under punkt (2) vil det konkrete undervisnings- og læringsmiljøet, slik studenter oppfatter og erfarer det, i vesentlig grad påvirke hvordan de orienterer og nærmer seg studie- og læringsarbeidet, dvs. hvilke arbeidsmåter og studiestrategier de velger å bruke. Noen undervisnings- og læringsmiljøer støtter, fremmer og bidrar til utvikling av effektive strategier og arbeidsmåter, andre gjør det ikke. Studenters opplevelse (perception) av og erfaringer med undervisning og læringsmiljø framstår så å si som en ”joker” i dette samspillet. Men det handler ikke om undervisning og studieprogrammer per se, som ”objektive” fenomener. Derfor er det avgjørende å få tilgang til *studenters erfaringer* og opplevelser. I neste omgang danner det grunnlag for å kunne si noe mer presist om og muligens predikere mer eksakt hvilke læringsstrategier og -mønstre studentene høyst sannsynlig vil engasjere seg i, og utvikle i det aktuelle undervisnings- og læringsmiljøet – om dette er læringsmønstre og arbeidsmetoder som fører dem fram til relativ suksess, eller fiasko i det akademiske læringsarbeidet.

Den skisserte forståelsesmodellen er tydelig og uttalt inspirert av forskningen på *læringstilnærminger* (approaches to learning) ved Universitetet i Gøteborg på 1970-tallet (jf. Entwistle 1987, 1988, 1997, 2000). Her ble begrepene *overflatetilnærming* (surface approach) og *dybdetilnærming* (deep approach) introdusert av den såkalte INOM-gruppa (’In-lärning och omvärdsuppfatning’) med Ference Marton, Roger Säljö, Lars Owe Dahlgren og Lennart Svensson som sentrale aktører (se Marton & Säljö 1976; Marton, Hounsell & Entwistle 1984; Marton et al. 1977).

Begrepsparet overflate- og dybdetilnærming beskriver kvalitativt ulike måter å handtere læringsoppgaver på i akademiske kontekster, og det representerer en milepæl i forskning på studenters læring i høyere utdanning de siste 25 årene (Webb 1997); eksempelvis er Marton og Säljös (1976) artikkel *On qualitative differences in learning – I: On outcome and process*, en av de mest siterte referansene i pedagogisk psykologi og læringsforskning innen høyere utdanning.

Typiske kjennetegn på de *kvalitative forskjellene* som begrepene -’deep’ og ’surface approach’ henspiller på, er sammenfattet i stikkordsform i tabell 1.

Et sentralt poeng ved *approaches to learning*, som også reflekterer det grunnleggende perspektivet på læring i Gøteborggruppas forskning, er at ”... people act accordingly to their interpretation of a situation rather than to «objective reality»” – presisert slik: ”The ways students learn are a function of how they perceive the learning task and the learning environment” (Watkins 2001:167). Læringstilnærminger (*approaches*) er *ikke* primært å forstå som relativt stabile, personforankrede og situasjonsuavhengige egenskaper eller karakteristikk. I denne forstand er lærings-tilnærming forskjellig fra det som gjerne omtales som *kognitiv stil* eller *læringsstil*, som viser til relativt stabile, individuelle kognitive *preferanser* og *særpre*g den enkelte mobiliserer på tvers av ulike oppgaver og situasjoner (jf. Riding & Rayner 1998; Schmeck 1988; Sternberg & Zhang 2001).

Tabell 1

Dybdetilnærming	Overflatetilnærming
Studentene:	Studentene:
<ul style="list-style-type: none"> – fokuserer på tekstens innhold, mening og budskap som en helhet – oppfatter forståelse av kunnskapsstoff, ideer og teorier som et mål i seg selv – forsøker å organisere og strukturere lærestoffets ulike deler til et større sammenhengende hele – relaterer tidligere kunnskap til ny kunnskap for å etablere ny forståelse og nye perspektiver – kobler teoretiske perspektiver til hverdagens erfaringer – og forsøker å gjøre sider ved virkeligheten mer forståelig og ”synlig” i lys av lærestoffets begreper og modeller. 	<ul style="list-style-type: none"> – fokuserer på tekstens detaljer: fakta, informasjon og begreper – mer eller mindre løsrevet fra meningen og det totale budskapet – forholder seg gjerne ”minimalistisk” til studiearbeidet – prøver å komme seg gjennom studier og kurs med minst mulig innsats – forbinder fakta og begreper på en ureflektert måte, som gjerne medfører reduksjon av lærestoffets meningsinnhold – lærer nytt stoff primært for å kunne reproducere det: memorerer og puffer med tanke på prøver og eksamen – oppfatter lærestoffet som pålagt ”utenfra”, som prøvekrav og kunnskap ”frikoblet” og isolert fra hverdags erfaringer og det virkelige liv.

Læringstilnærminger beskriver primært den enkelte students *intensjoner* med og motiver for læringsarbeid: *Det* den enkelte har til hensikt å oppnå ut fra erfaringer med bestemte læringssituasjoner og oppgaver; læringstilnærminger er "... adopted in response to the manner in which a particular course or curriculum program is delivered" (Richardson 1990:165). Begrepet læringstilnærming relateres til en kontekstuell og relasjonell forklaringsmodell, nært beslektet med sosiokulturelle og situerte perspektiv på læring – og: "... consistent with the current advocacy of situated cognition" (Watkins op cit:166). Det markerer en vending vekk fra et mer indidorientert forskningsperspektiv på studenters læring, både når det gjelder læreprosesser, læringsresultater og akademiske prestasjoner.

Kort oppsummert har studenters grunnleggende tilnærming (approach) til læring betydelig effekt på deres læringsresultater (learning outcomes), det er Marton og Säljös (1976) sentrale poeng: Studentenes oppfatning av de krav læringsoppgaver og -situasjoner stiller, representerer en avgjørende forklaringsvariabel. Opplevd undervisningskvalitet, i vid forstand, avgjør om studenter adopterer en utpreget *dybdetilnærming*, som gjennomgående resulterer i en grunnleggende forståelse og innsikt i lærestoffet. Alternativet er at de anlegger en typisk *overflatetilnærming*, som fører til mer fragmentariske læringsresultater preget av manglende forståelse og meningsetablering, som generelt vises som svake akademiske prestasjoner.

"The Course Perception Questionnaire" (CPQ) ble utviklet innenfor perspektivet skissert over for å kartlegge variasjoner i studenters oppfatninger av ulike akademiske kontekster og læringsmiljøer i større studentpopulasjoner. Gøteborgstudiene var i hovedsak basert på kvalitative og såkalte fenomenografiske metoder med mindre studentgrupper (typisk, n=30–40). I den andre Lancasterstudien ønsket man i tillegg å belyse problemstillingene med *kvantitative* metoder, med standardiserte spørreskjemaer og inventories. Både CPQ og *Approaches to Studying Inventory* (ASI) – utviklet for å kartlegge studenters læringsmønstre (bl. a. dybde- og overflatetilnærming) – inngår i denne Lancasterstudien, og representerer i en viss forstand "a questionnaire variant of Marton's procedure" (jf. Entwistle & Ramsden 1983; Pettersen 2004, 2006; Richardson 2000).

Spørsmålene i den første versjonen av *Course Perception Questionnaire* var basert på et omfattende intervjumateriale med studenter i ulike fag- og utdanningskontekster. Denne versjonen ble testet i en pilotstudie med et utvalg studenter som deretter ble intervjuet om sine erfaringer med spørreskjemaet. Den endelige versjonen av CPQ består av 40 spørsmål fordelt på følgende *åtte* skalaer/faktorer (hver faktorskåre er basert på fem spørsmål) (Entwistle & Ramsden 1983:121–130):

- | | |
|------------------------------|-------------------------|
| 1. Formal teaching methods | 5. Good teaching |
| 2. Clear goals and standards | 6. Freedom in learning |
| 3. Workload | 7. Openness to students |
| 4. Vocational Relevance | 8. Social climate |

Fra "course perception" til "course experience"

Resultater fra den andre Lancasterundersøkelsen er dokumentert i *Understanding Student Learning* (Entwistle & Ramsden 1983). Prosjektets hovedtema er det samme som i Gøteborgstudiene: "To demonstrate the connetions between students' approaches and the context of learning" (ibid:198) – men basert på vesentlig større utvalg: Totalt 2.208 studenter fra 54 ulike institusjoner i Storbritannia deltok. Prosjektets hovedkonklusjoner basert på så vel kvalitative som kvantitative metoder oppsummerer Entwistle og Ramsden slik:

... combining the impression of causality from the interviews with the evidence of relationship from the questionnaires, we can begin to piece together a chain of causality ...

Positive attitudes to studying, deep approach, and academic progress are all related to *good teaching, freedom in learning, and avoidance of overloading*. ... Combined with assessment methods perceived to be appropriate, *these contextual characteristics increase the probability that students will take deep approaches* (ibid:202; min kursivering).

Og på den andre siden:

... perceptions of inappropriate or excessive assessment, together with a too rigidly structured curriculum, encourage extrinsic motivation, engender poor attitudes, and thereby make surface approaches more likely (ibid:203; min kursivering)

Sterkest samvariasjon (korrelasjon) mellom studenters læringstilnærminger og opplevd undervisningskvalitet basert på ASI og CPQ gjelder i hovedsak disse forholdene (jf. Richardson 2000:89ff):

- Høye skårer på læringsmønsteret *meningsorientering* (som inkluderer dybdetilnærming) korrelerer signifikant med god undervisning og studentautonomi (freedom in learning).
- Høye skårer på læringsmønsteret *reproduserende orientering* (som inkluderer overflatetilnærming) korrelerer signifikant med lav studentautonomi og høy opplevd arbeidsbyrde (workload). Sterkest korrelasjon totalt sett er mellom reproduserende orientering/overflatetilnærming og høy opplevd arbeidsbyrde. Dette forholdet er dokumentert i flere replikasjonsstudier, blant annet i en større studie av Meyer og Parsons (1989) med to utvalg fra et sørafrikansk universitet.

Når Lancasterstudien ikke dokumenterer flere og klarere statistiske sammenhenger, forklares det gjerne med at det hefter psykometriske svakheter ved *Course Perception Ouestionnaire*. I noen grad gjelder det også den originale versjonen av *Approaches to Studying Inventory* (jf. Lizzio, Wilson & Simons 2002; Richardson 2000, 2005). Paul Ramsden arbeidet med å videreutvikle spørreskjemaet, basert på erfaringene med CPQ. Han støttet seg også på andre tilsvarende evalueringsverktøy, blant annet *School Experience Questionnaire* utviklet for elever i grunnskolen (Ramsden et al. 1989) og Entwistle og Taits (1990) *Experiences of Studying and Higher Education Questionnaire*. Dette arbeidet resulterte i *The Course Experience Questionnaire* (CEQ) (Ramsden 1991).

Den såkalte Linke-komiteen (Performance Indicators Research Group, PIRG) anbefalte å teste CEQ som "... an appropriate indicator of perceived teaching quality" – som en mulig "performance indicator" (PI) for undervisningskvalitet og -effektivitet i det nasjonale kvalitetssystemet for høyere utdanning i Australia. En forutsetning var at spørreskjemaet

hadde maksimum 30 item. Tanken var at CEQ kunne inngå som element i den årlige studentundersøkelsen, *The Graduate Destination Survey* (GDS), som sendes alle avgangstudenter i Australia omtrent fire måneder etter avsluttet studium (Linke 1991).

Course Experience Questionnaire med 30 item ble prøvd ut i 1989 med til sammen ca. 4.500 studenter fra 50 ulike institusjoner. Denne 30-itemsversjonen, anbefalt som element i GDS, har følgende fem skalaer; de fire første er videreført fra *Course Perception Questionnaire* (Ramsden 1991):

- (1) *Good teaching* (utvidet til 8 item).
- (2) *Clear goals* (5 item).
- (3) *(Appropriate) Workload* (5 item).
- (4) *Emphasis on Independence* (Freedom in Learning, utvidet til 6 item).
- (5) *Appropriate Assessment* (ny skala, 6 item)

CEQ har samme intensjon som CPQ: Kartlegging av forhold som studenter erfarer som betydningsfulle og kritiske i forhold til utviklingen av læringsproduktive studiestrategier og effektiv studieatferd. De fem skalene som utgjør kjernen i CEQs teoretiske referanseramme, er definert og operasjonalisert slik (jf. Lizzio et al. 2002; se også Ramsden 1992:Kap 6):

- *Good teaching* (God undervisning) – viser til hvorvidt det akademiske undervisnings- og læringsmiljøet oppfattes som *engasjerende* og *involverende*. Det kommer bl. a. til uttrykk ved at lærerne:
 - viser ekte interesse for studentenes synspunkter og meninger;
 - forsøker å forstå vansker som studenter måtte ha og møter i sitt studiearbeid;
 - kommuniserer positive forventninger og søker å motivere studenter til å yte sitt beste i studiearbeidet;
 - gir gode forklaringer på ideer, teorier og begreper som formidles i undervisningen;
 - bruker tid og krefter på å gjøre lærestoffet/fagene interessant;
 - gir studentene nyttige tilbakemeldinger (feedback) om framgang og hvordan de ligger an faglig.

De forhold som særlig løftes fram i denne operasjonaliseringen av god undervisning, omtales også som de 3 E'ene: *Enthusiasm, Explanation, Empathy*.

- ◉ *Clear goals* (Klare mål) – viser til læringsmiljøer der man vet hva lærerstaben forventer av studentene og deres prestasjoner; dette kommer til uttrykk ved at lærerne:
 - ”helt fra begynnelsen” forklarer hvilke læremål som gjelder og vektlegges i kurset/utdanningen/faget;
 - markerer tydelig progresjonen i undervisning og studieopplegg; viser hvor ”vi er” og ”hvor vi er på veg”;
 - forklarer eksplisitt standarder og kriterier som studenters læringsarbeid og -resultater blir vurdert etter.
- ◉ *Appropriate workload* (Passe, overkommelig arbeidsbyrde) sikter til undervisningskontekster som bidrar til å regulere studentenes arbeidsbyrde og -belastning på en effektiv måte, ved at:
 - det er praktisk mulig å komme gjennom og tilegne seg lærestoff som inngår i læreplaner og dekkes i undervisningen;
 - bredden i læreplanens temaer og emner er tydelig presisert og avgrenset;
 - studieopplegg og undervisning sørger for at det gis nok tid til å fullføre studieoppgaver og arbeidskrav (jf. ”time on task”).
- ◉ *Appropriate assessment* (Tilpassede, relevante vurderingsformer) referer til vurderings- og eksamensformer som fremmer og støtter læringsproduktive prosesser, ved at de:
 - legger vekt på forståelse og selvstendig tenkning, det legges liten vekt på memorering og pugg av faktakunnskaper;
 - er tilpasset studiets progresjon og jevnt fordelt i studieforløpet;
 - omfatter (formativ) feedback utover karaktersetting og korte, skriftlige tilbakemeldinger på oppgaver og innleveringsarbeider.

- *Independence in learning* (Studentautonomi, valgfrihet) – viser til læringskontekster hvor studentene har en viss valgfrihet (innenfor gitte rammer), når det gjelder:
 - utvikling av akademiske interesser og muligheter for å forfølge egne faglige interesser i studiet;
 - hvordan studenter arbeider med og tilegner seg lærestoffet (studie-strategier og arbeidsmåter);
 - måten studiearbeid og oppgaver blir vurdert på (vurderings- og eksamensformer).

I tillegg til de 30 spørsmålene, fordelt på fem skalaer, inneholder CEQ også et spørsmål som gjelder studentenes *tilfredshet/fornøydhet* med kvaliteten på kurset (studieprogrammet): "Overall, I was satisfied with the quality of this course".

Generelle ferdigheter og kompetanseinnretninger

Utviklingen av en ny, sjette skala, *Generic Skills Scale* (GSS), er resultat av samarbeidet mellom *The Department of Employment, Education and Training* (DEET) og *The Graduate Careers Council of Australia* (GCCA), som administrerer det nasjonale kvalitetssikringssystemet i Australia. Samarbeidet omfatter også arrangement av konferanser om erfaringene med CEQ som kvalitetsindikator og utvikling av CEQ som evaluerings-instrument, såkalte *Course Experience Symposium* (jf. Hand & Trembart 1999; McInnis et al. 2001).

Skalaen *Generelle ferdigheter* reflekterer en økende interesse for hvordan ulike utdanningskontekster og læringsmiljøer påvirker utviklingen av såkalte *generic skills* eller *personal, transferable competencies*. Kort skissert dreier det seg om: "... the process skills which help students to effectively apply the content or subject skills learnt in higher education to work environment ... variously termed *key skills, core competencies or generic attributes* ... [which] include problem solving, analytic skills,

teamwork and the ability plan work” (Lizzio et al. 2002:31; min kursivering; se også Bowden & Marton 2004).

The Generic Skills Scale (GSS) består av følgende item/utsagn:

- This course has helped me to develop my problem–solving skills.
- This course has sharpened my analytic skills.
- This course has helped develop my ability to work as a team member.
- As a result of doing this course, I feel more confident about tackling unfamiliar problems.
- This course has improved my written communications skills.
- This course has helped me develop my ability to plan my own work.

Ulike versjoner av CEQ

Kombinasjoner av Ramsdens originale versjon av CEQ30 og skalaen *Generic Skills Scale* danner grunnlag for ulike varianter av CEQ. To standardversjoner omtales henholdsvis som den *lange* og *korte* versjonen: Den lange versjonen, eller CEQ36, består av CEQ30 pluss GSS; i alt 37 item, spørsmålet om generell tilfredshet med kurset/studiet medregnet (jf. Wilson, Lizzio & Ramsden 1997; Richardson 2005).

Den korte versjonen er en *modifisert* versjon av CEQ – CEQ23 – som har vært et fast element i Graduate Destination Survey i Australia siden 1993. CEQ23 er modifisert på følgende måte: (1) *Generic Skills Scale* erstatter skalaen *Independence in learning* (Studentautonomi); begrunnelsen for å utelate denne skalaen er dens relativt svake begrepsmessige validitet og skalastruktur; (2) de øvrige fire skalene er forkortet ved at de psykometrisk sett svakeste itemene er utelatt; bl.a. basert på resultater fra Ramsdens (1991) pilotstudie. Det gir følgende fordeling av item på de fem skalaene i CEQ23:

God undervisning	6 item	Arbeidsbyrde	4 item
Generelle ferdigheter	6 item	Vurderingsformer	3 item
Klare mål	4 item		
=			23 item
+	spørsmål 24, om generell tilfredshet med kursets kvalitet		

En alternativ kortversjon – *CEQ24* – har et ekstra item knyttet til skalaen *Vurderingsformer*: ”The assessment methods employed in this course required an in-dept understanding of the course content”. Det viser seg imidlertid at utsagnet primært oppfattes som et aspekt ved god undervisning, basert på faktoranalyse av et større utvalg bachelorgradsstudenter ($n \approx 58.000$) (Ainley 2001). Det er de to kortversjonene – *CEQ23* og *CEQ24* – som er hyppigst anvendt i ulike utdanningskontekster, også utenfor Australia.

EMS, norske varianter av CEQ

De tre norske variantene av *CEQ*² som danner grunnlag for valideringsstudiene i rapportens del 3, omtales som *Erfaringer med studiet*, forkortet *EMS*. De seks skalaene som inngår i de ulike variantene, er gitt følgende betegnelser og forkortelser:

- | | |
|--|--|
| • GU God undervisning | • ARB Tilpasset arbeidsbyrde |
| • GF Generelle ferdigheter
(kompetanseinnretninger) | • SA Studentautonomi (valgfrihet i
læringsarbeidet) |
| • KM Klare mål | • VF Relevante vurderingsformer |

De tre *EMS*–variantene har følgende skalastruktur:

- *EMS30* er direkte oversatt fra Ramsdens *CEQ30*, og har samme struktur.
- *EMS30–GF* bygger på *CEQ30*, men skalaen *SA* er erstattet med *GF*.
- *EMS26* er direkte oversatt fra *CEQ23*, men utvidet med tre item fra skalaen *SA*. *EMS26* omfatter dermed alle seks skalaer, akkurat som *CEQ36*.

² Samtlige versjoner er oversatt av forfatteren.

DEL 2

Studentevaluering og CEQ/EMS i praksis

Prinsipper, strategier og eksempler

Studentevaluering – intensjoner og strategier

Studentevaluering av undervisning kan sammenfattende beskrives som forskjellige former for tilbakemelding fra studenter om deres erfaringer med og vurdering av ulike sider ved undervisning og læringsmiljø. Evalueringsstrategier kan ha ulike intensjoner og adressater. CEQ, som standardisert spørreskjema for evaluering av opplevd undervisningskvalitet, representerer én strategi og inngang for å sikre systematisk tilbakemelding fra studenter.

I en større oversiktsartikkel skisserer Marsh (1987:259) hvordan studentevaluering sikter på å skaffe data til ulike formål, hvor de vanligste er:

1. Diagnostic feedback to faculty about the effectiveness of their teaching that will be useful for the improvement of teaching.
2. A measure of the teaching effectiveness to be used in administrative decision making.
3. Information for students to use in the selection of courses and instructors.
4. A measure of the quality of the course, to be used in course improvement and curriculum development.
5. An outcome or a process description for research on teaching.

Et typisk dilemma i studentevaluering er spørsmålet om datainnsamling for et formål også er relevant for andre formål; det gjelder særlig "... whether the same data should be used for diagnostic feedback to faculty and for administrative decision making" (ibid:260). Dette berører et vanlig skille mellom evaluering med et *formativt*, utviklingsorientert siktepunkt og evaluering med en *summativ*, kvalitetskontrollerende og kvalitetssikrende intensjon (jf. Bowden & Marton 2004; Handal 1996, 2006; Lauvås & Jakobsen 2002; Pettersen 2005).

Formative og summative intensjoner

Formativ, utviklingsorientert studentevaluering knyttes gjerne til evaluering på undervisnings- og kursnivå, som en systematisk tilrettelagt dialog mellom studenter og lærere som er *direkte involvert* i undervisningen. Denne formen forutsetter at lærer(e) reelt ønsker å justere, utvikle og forbedre aspekter ved undervisnings- og læringsprosessen som det er realistisk og mulig å gjøre noe med, i samarbeid med studentene. Eventuelle endringer og justeringer gjelder primært studentene som deltar i undervisningen som evalueres. I forhold til Marshs (1987) kategorisering handler det om ”diagnostic feedback to faculty”. Lærere har et vidt spekter av metoder til rådighet i utviklingsorientert evaluering, men ofte framheves kvalitative metoder som særlig egnet for denne evalueringsformen; metoder der dialog og direkte kommunikasjon mellom lærer og studenter står i sentrum (jf. Handal op.cit).

Studentevaluering med sikte på *kvalitetskontroll og -sikring* retter seg i større grad mot utdanningens ledelse; adressater er enkeltpersoner, grupper og/eller utvalg med ansvar for enhetens/institusjonens kvalitets-sikringssystem. Strategien har på den ene siden en uttalt summativ intensjon: hensikten er å dokumentere i hvilken grad undervisning og studieprogrammer innfrir krav og standarder som stilles til undervisnings- og programkvalitet (’accountability’). I forhold til Marshs kategorier dreier dette seg om ”measure to be used in administrative decision making”. På den andre siden er tanken at evalueringresultatene også skal fungere som informerte handlingsdata for hva som kan /bør/skal gjøres, dersom vedtatte kvalitetsstandarder ikke innfris. Med andre ord inkluderes formative intensjoner: Hensikten er at studentevalueringer også kan anvendes ”in course improvement and curriculum development” (jf. Marsh op cit).

Evaluering for *kvalitetspresentasjon* er et *tredje* formål med studentevaluering. Offentlighet og potensielle studenter er primære adressater; et forhold Marsh refererer til som “information for students to use in the selection of courses”. Med andre ord inngår studentevaluering som ledd i markedsføringsstrategier, som kan kobles til en økende oppmerksomhet på

studenter som ”kunder” og utdanningsinstitusjoner som ”leverandører” (jf. Handal 2006).

Kvantitative tilnærminger, f. eks ulike typer *spørreskjemaer*, er ofte førstevalget i studentevaluering for kvalitetskontroll og kvalitetspresentasjon. Kvantitative indikatorer, som frekvenser, prosenter, indeksverdier og faktorskårer er enkle å aggregere fra flere kurs og studieprogrammer med store studentgrupper. Denne typen kvantitative *kvalitetsindikatorer* (performance indicators) kan dermed danne grunnlag for å overvåke (monitor) og vurdere endringer i undervisningskvalitet *over tid* – og forskjeller *mellom* ulike kurs og studieprogrammer på et gitt tidspunkt.

Studentevaluering som indikasjon på undervisningskvalitet

Det finnes ingen enkel, omforent og endelig definisjon av god undervisning; like lite som det finnes enkle og entydige indikatorer for undervisningskvalitet. Flere indikatorer kan tenkes ved siden av studentevaluering av undervisning: Læreres selvevaluering av egen undervisningskompetanse og kollegavurdering av læreres undervisningskompetanse, er to eksempler. Når studentevaluering har fått en sentral rolle som indikator for undervisningskvalitet, er begrunnelsen blant annet at det har vist seg problematisk å finne samme empiriske støtte for andre indikatorer: ”The heavy reliance on SETs [students’ evaluations of teaching] as the primary measure of teaching effectiveness stems in part from the lack of support for the validity of any other indicator of effective teaching” (Marsh & Roche 1997:1190).

Det er også godt forskningsmessig belegg for at ”... SETs reflects students’ learning”, et forhold som er konsistent med SAL-perspektivet. Her framheves også studenters opplevelse av undervisning og læringsmiljø som forklaringsfaktor i forhold til de læringsresultater og akademiske prestasjoner studenter oppnår (jf. del 1). Flere studier og meta-analyser anslår at korrelasjonen mellom undervisningsevaluering og studenters prestasjoner ligger rundt 0,30–40, mens korrelasjonen i forhold til eksamenskarakterer gjennomgående ligger mellom 0,20 og 0,30 (Marsh & Roche op.cit; se også Marsh 1987).

Derneft, i et utviklingsorientert perspektiv, viser det seg at lærere som får feedback og konsultasjon på egen undervisning basert på studentevalueringer, gjennomgående får bedre evalueringer senere – og: ”SET feedback and consultation are effective means to improve teaching effectiveness” (Marsh & Roche 1997:1194).

Oppsummert underbygger hovedtyngden av senere forskning de konklusjoner og synspunkter H. H. Remmers presenterte på 1950-tallet. Remmers initierte systematiske forskningsprogrammer på studentevaluering allerede på slutten av 1920-tallet og regnes som en pioner på feltet. Et utvalg av Remmers’ konklusjoner og synspunkter er framhevet nedenfor (etter Marsh 1987:258f):

- Studentevalueringer er valide uttrykk for undervisningskvalitet ”also measured by what students actually learn of the content of the course”.
- “Teachers at all levels of the educational ladder have no real choice as to whether they will be judged by those whom they teach ... The only real choice any teacher has is whether he wants to know what these judgement are and whether he wants to use this knowledge in his teaching procedures”.
- “... students are pretty much the only ones who observe and are in a position to judge the teacher’s effectiveness”.
- “Knowledge of student opinions and attitudes leads to improvement of the teacher’s personality and educational procedures”.
- “No research has been published invalidating the use of student’s opinion as one criterion of teacher effectiveness”.

CEQ som utviklingsorientert evalueringsverktøy

Course Experience Questionnaire kan anvendes både formativt, som utviklingsorientert evalueringsverktøy, og som element i summative kvalitets-systemer: "... the CEQ is not just an indicator of educational quality. It also points to possibilities for educational improvement" (Patrick 2003:7). I det australske kvalitetssystemet går det imidlertid relativ lang tid fra evalueringen (Graduate Destination Survey) til resultatene foreligger; noe som svekker CEQs utviklingsorienterte potensial. For å bøte på dette har University of Sydney utviklet *Student Course Experience Questionnaire* (SCEQ), hvor CEQ24 er inkludert (Ginns 2003). SCEQ med totalt 67 item administreres årlig til alle studenter som ledd i universitetets kvalitetsprogram (første gang i 1999). Dermed får lærerne raskere tilgang til CEQ-data som utgangspunkt for videre evaluering av egne kurs. Det gir bedre muligheter for at CEQ og lærernes egne evalueringsstrategier kan fungere *komplementært* (jf. Pettersen 2005:kap 14). Generelt innbærer det å utnytte fordelene ved varierte metodiske tilnærminger til evaluering av undervisning og læringsmiljø (jf. metodetriangulering). Entwistle (2005) løfter fram dette forholdet slik: "... the interplay between questionnaire and interview data, and *the value of having discussions with groups of students in considering areas of difficult and possible improvement* cannot be over-emphasised" (min kursivering).³

Presentasjon og analyse av CEQ-data

Lærere og administratorer som ønsker å bruke CEQ bør være fortrolige med et knippe praktiske og teoretiske grep: i hovedsak samme forhold som gjelder bruk av standardiserte spørreskjemaer generelt. To momenter skis-

³ Hentet fra Noel Entwistles oppsummering av erfaringer fra prosjektet *Enhancing Teaching–Learning Environments in Undergraduate Courses* (2000–2004), The ETL project, se: www.ed.ac.uk/etl

seres nedenfor: (1) prosedyre for administrasjon og skåring av CEQ; (2) vanlig praksis for presentasjon, analyse og tolkning av CEQ-data.

Alle CEQ-/EMS-versjoner anvender en *fempunktsskala* (Likert) der studentene tar stilling til alle utsagn/spørsmål ut fra følgende alternativer:

- 5= Helt enig
- 4= Delvis enig
- 3= Nøytral (Usikker/Ikke relevant/Kan ikke svare)
- 2=Delvis uenig
- 1=Helt uenig.

Flere utsagn i CEQ er formulert *negativt*, for eksempel: ”For å klare seg godt på dette studiet er alt du egentlig trenger en god hukommelse” (Vurderingsformer, VF). Skåres dette utsagnet ”Helt enig” (5), bidrar det til å *høyne* samlet skåre på denne skalaen; det er imidlertid *motsatt* av operasjonaliseringen av *Vurderingsformer* i CEQ. For å standardisere skala-/faktorskårer slik at høye verdier reflekterer positive erfaringer, må alle ”negativt” formulerte utsagn *inverteres før* videre bearbeiding av rådata (etter formelen: $6 \div \text{råskåre}$). Vedlegg 1–3 inneholder en oversikt over negativ formulerte utsagn i de ulike EMS-versjonene (se også tabell 4, s. 50).

Faktorskårer

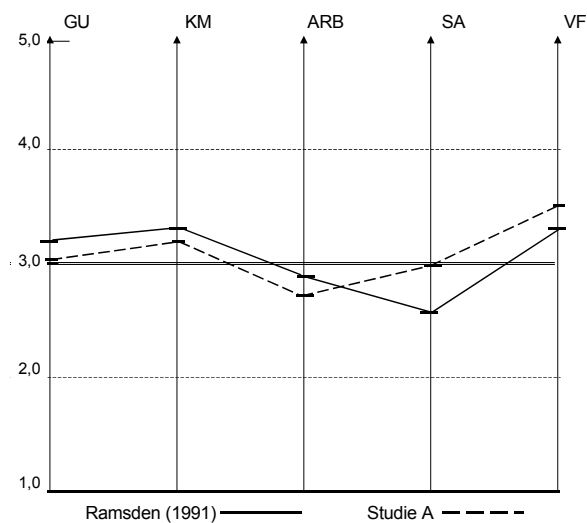
Faktorskårer er gjennomsnittsverdier for utsagnskårer i de respektive faktorene (skalaene);⁴ i eksemplene nedenfor anvendes norske forkortelser for skalaene (se s. 24). Presentasjon av faktorskårer i tabells og/eller grafisk form er en vanlig måte å formidle CEQ/EMS-resultater på. Eksemplene i figur 1 og tabell 2 viser resultater fra Ramsdens (1991) studie med CEQ30 og resultater basert på EMS30 med 35 kull/klasser i norske helse- og sosialfagutdanninger (fra studie A, Del 3).

⁴ I CEQ/EMS anvendes *uvektede* faktorskårer: Utsagn (variabler) som inngår i en faktor/skala tillegges lik vekt (jf. Ullenberg & Nordvik 2001).

Tabell 2

		CEQ-/EMS-faktorskårer				
		GU	KM	ARB	SA	VF
CEQ30	Ramsden (1991), n=3.372	3,2	3,3	2,9	2,6	3,3
EMS30	Studie A, n=1.611	3,0	3,1	2,7	3,0	3,5

Grafisk presentasjon av faktorskårer



Figur 1

Faktorskårer basert på aggregerte datasett fra ulike og til dels svært forskjellige undervisningskontekster fungerer først og fremst som referanseverdier for studentpopulasjoners vurdering av opplevd undervisningskvalitet. Aggregerte data gir også et bilde av forholdet mellom enkeltfaktorene, såkalte "CEQ-profiler". Gjennomsnittsverdiene for God Undervisning, Klare mål og Vurderingsformer i figur 1 ligger nær, men noe over *nøytral* verdi (3,0) i begge utvalgene, mens Arbeidsbyrde og Studentautonomi skåres lavere, under 3,0. Dette er en relativt vanlig CEQ-profil for aggregerte datasett.

Når resultater brytes ned på utdannings-/studieprogramnivå, er variasjonen i faktorverdiene vanligvis større; aggregering maskerer i en viss forstand dette ved at skårene tenderer mot middelveien 3,0. Derfor er også sammenlikning av opplevd undervisningskvalitet på institusjonsnivå ut fra aggregerte resultater lite fruktbart: "... aggregating CEQ scores within a university does not produce meaningful result, comparing aggregated sco-

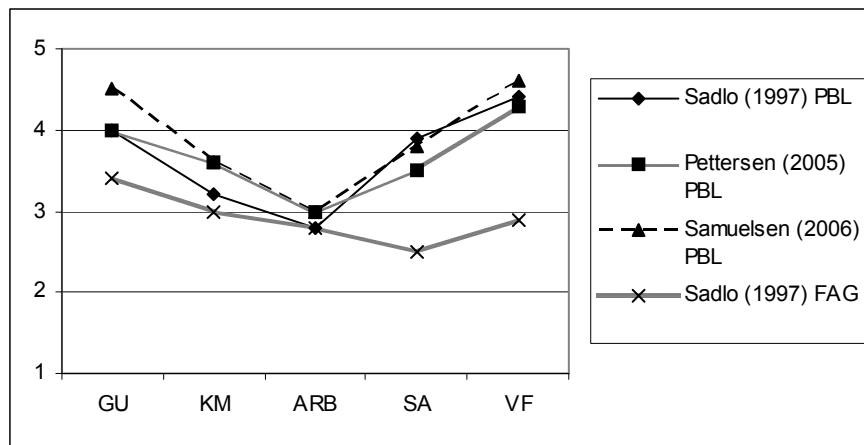
res between universities is pointless (Hand & Trebarth 1999:58). Komparative studier bør primært ta utgangspunkt i resultater på kurs- og programnivå. Både når det gjelder sammenlikninger over tid for samme kurs, og mellom samme eller tilsvarende kurs/programmer som tilbys ved ulike institusjoner. I begge tilfeller er det viktig å kjenne responsrate for de kursene det gjelder; slik at vi bedre kan vurdere hvor representative de aktuelle evalueringsresultatene er.

Med forbeholdene over i mente er det *Studentautonomi* som avviker mest i de to utvalgene (figur 1), med en verdi i studie A som er 0,4 høyere sammenliknet med Ramsdens studie. En mulig forklaring kan være at samtlige studieprogrammer i studie A bygger på problembaserte læringsprinsipper, om enn i varierende grad. En studie av Sadlo (1997) som sammenlikner problembaserte og fagbaserte ergoterapiutdanninger basert på CEQ-resultater, rapporterer enda større differanser for skalaen Studentautonomi (jf. Sadlo & Richardson 2005, se figur 2 nedenfor). De øvrige forskjellene i tabell 2 er på 0,2, og dermed er av mindre interesse; det er primært signifikante differanser, *minimum* 0,4–0,5 (ca. et halvt standardavvik), som bør fange oppmerksomheten med tanke på nærmere oppfølging og analyse.

Figur 2 viser også et eksempel på *grafisk* datapresentasjon. Faktorskårene er fra evalueringer på *kull-/klassenivå* for fire helse- og sosialfaglige utdanningsprogrammer: (1) To utvalg fra Sadlos studie (op. cit.), det problembaserte og fagbaserte programmet med hhv. høyeste og laveste faktorskårer; (2) en spesialistutdanning i sykepleie (Samuelsen 2006); (3) en grunnutdanning i vernepleie (Pettersen 2005).⁵ Spesialistutdanningen og utdanningen i vernepleie representerer også problembaserte studieprogrammer.

De problembaserte programmene har generelt høye skårer, med Arbeidsbyrde som et klart unntak, og faktorprofilene har tilnærmet samme form. Forskjeller av interesse *mellom* PBL-utdanningene gjelder God undervisning og Klare mål, hvor differansen mellom to av programmene er $\geq 0,4$.

⁵ Upublisert studie gjennomført i 2000–02; data fra studien inngår imidlertid i Studie C (jf. Del 3), og studien er omtalt nærmere i Pettersen (2005).



Figur 2

Det bør også være av interesse å se nærmere på den relativt sett lave skåren på Arbeidsbyrde, siden det er en uttalt intensjon i PBL å stimulere og støtte forståelsespreget og dybdeorientert læring, mens lave skårer på Arbeidsbyrde gjerne assosieres med læringsmiljøer hvor studenter støtter seg på memorering og overflatepregede studiestrategier.

Imidlertid gir ikke faktor- eller skalaskårer i seg selv umiddelbare forklaringer på dette tilsynelatende paradokset; de gir i første omgang kun en *generell* indikasjon på opplevd undervisningskvalitet; "... the CEQ data provides [only] a first layer of information" (Bowden & Marton 2004). Det gjelder også differansene i forhold til det fagbaserte programmet (Sadlo 1997) med en CEQ-profil som ligger nærmere den generelle profilen antydnet tidligere, med verdier under gjennomsnittet både for Studentautonomi og Arbeidsbyrde. I tillegg ligger verdien for God undervisning signifikant under samtlige PBL-programmer.

Et forhold som er viet oppmerksomhet mht. presentasjon av CEQ-resultater, er hvordan dette kan gjøres på enklere, mer lettfattelig og "... more user-friendly at least for those users who have little understanding of mean scores, standards deviations ... and Cronbachs alpha" (Gallager 1999:22). En alternativ presentasjonsform for faktorskårer, som anvendes i australske CEQ-rapporter, går ut på å rekode skalaverdiene på følgende måte: 5= 100; 4= 50; 3= 0; 2= ÷50; 1= ÷100. Med andre ord settes skalagjennom-

snitt til ”0”; skårer over gjennomsnittet får positive verdier, mens verdier under gjennomsnittet tildeles negative verdier. Endringer/forskjeller av interesse er minimum 25 poeng, dvs. et halvt standardavvik (jf. Hand & Trebarth 1999).

Tabell 3 viser et eksempel på rekodede faktorskårer; resultatene er fra bachelorutdanningene i *Graduate Destination Survey 2000* (Ainley 2001) og et utvalg norske høgstskolestudenter (fra studie C; jf. Del 3); tilfredshet med studiekvalitet er tatt med i tabellen (spm 24).

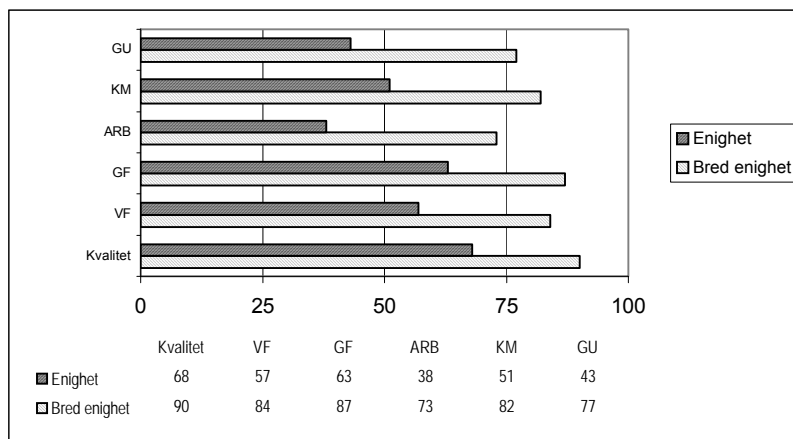
Tabell 3

		Rekodede faktorskårer					Kvalitet
		GU	KM	ARB	VF	GF	
CEQ23	Ainley (2001), n≈58.000	13	21	13	36	35	43
EMS26	Utvalg, studie C, n= 437	÷ 2	5	÷32	34	48	37

Prosentvis enighet – frekvensfordeling

Prosentvis enighet er en alternativ og utfyllende presentasjonsform som gir mer detaljert informasjon om responsmønstrene bak faktorskårene. Prosentvis enighet er summen av prosentvis frekvens for kategoriene *Helt enig* (5) og *Delvis enig* (4). I tillegg anvendes ofte prosentvis *bred enighet* som inkluderer kategorien *Usikker/Nøytral* (3). Når begge former for frekvensfordeling presenteres, gir det et oversiktsbilde av forholdet mellom tre hovedkategorier: uttalt *positive* skårer, *nøytrale* skårer og uttalt *negative* skårer. Figur 3 viser et eksempel på presentasjon av 'enighet' og 'bred enighet' på *faktornivå* (beregnet gjennomsnitt for hhv. enighet og bred enighet for utsagn som inngår i en skala/faktor). Resultatene i figur 3 er basert på data fra Ainley (2001), jf. tabell 3.

Presentasjon av prosentvis *enighet* og *bred enighet*



Figur 3

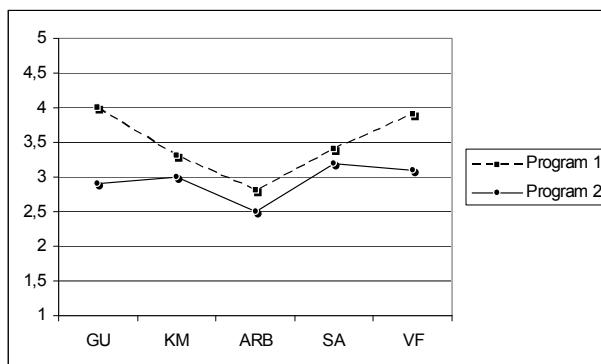
Presentasjon og analyse på program-/klassenivå – eksempler

Relevant analysenivå for å vurdere status, forskjeller og utviklingstrekk basert på CEQ-resultater er primært studieprogram- og/eller klassenivå (jf. Marsh 1987; Marsh & Roche 1997). Nedenfor skisseres et forslag til standardprosedyre for CEQ anvendt i utviklingsorientert studentevaluering, med følgende trinn: (a) Inspeksjon og vurdering av CEQ-profiler; gjelder både komparativ analyse (mellom reelt sammenliknbare programmer) og analyse av utvikling (samme program over tid); (b) Vurdering av responsmønstrene prosentvis enighet og bred enighet for faktorer med signifikante differanser; (c) Vurdering av enkeltutsagn som mest uttalt bidrar til de aktuelle differansene, (d) Rapportering og diskusjon av analyseresultatene (inkluderer trinn a, b og c) med studentgruppa med sikte på å utvikle og forbedre aspekter ved undervisning og læringsmiljø som analysen setter søkelyset på.

Utvalgte momenter i strategien er eksemplifisert med utgangspunkt i EMS30-data for to studieprogram innen helse- og sosialfag. Utdanningene følger samme rammeplan og har i tillegg utformet svært like fag- og undervisningsplaner basert på PBL, som lærerkollegiene samarbeidet om på felles læreplanseminarer. De to programmene er lokalisert på to høyskoler. Punkt (d) i strategien, rapportering og drøfting med studentgruppene, ble

ikke gjennomført i dette eksemplet. Figur 4 viser EMS–profiler fra evalueringer gjennomført i vårsemesteret første studieår; for program 1 er n=34 med responsrate 99%, for program 2 er n=75 og responsrate 85%.

EMS–profil for to utdanninger

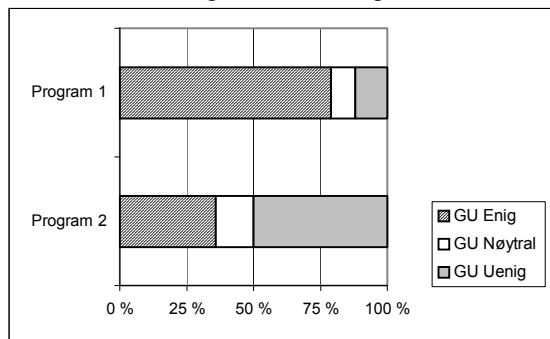


Figur 4

Figur 4 viser signifikante forskjeller mellom studieprogrammene for God undervisning (GU) og Vurderingsformer (VF).

Videre i eksempelet analyseres aspekter ved faktoren GU nærmere; for programmene er faktorskåre hhv. 4,0 og 2,9. Figur 5 viser faktorens frekvensmønster for *enighet*, *bred enighet* (enig + nøytral) og *uenighet* (helt + delvis uenig). Det er en relativ liten andel av studentene i program 1 som vurderer GU negativt (12%), mens halvparten av studentene i program 2 gir negativ vurdering av denne faktoren.

GU – Program 1 vs Program 2



Figur 5

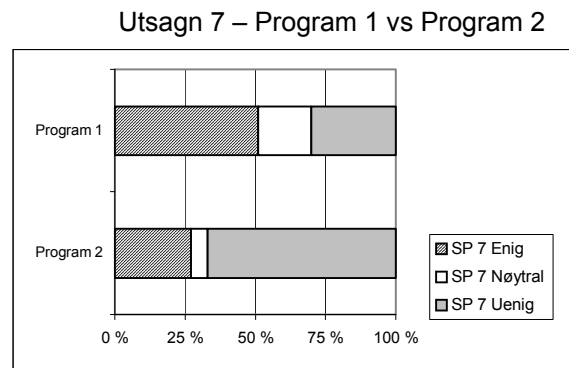
Neste steg er å analysere utsagn med lavere gjennomsnittverdi enn gjennomsnitt for faktorskåren. Et eksempel i begge utdanninger er utsagn 7: ”Lærerne bruker mye tid på å kommentere og gi tilbakemeldinger på studiearbeidene mine”, som gjelder *mengden* feedback fra lærerstaben som aspekt ved god undervisning. Figur 6 viser at det er markert forskjell i pro-

sentvis enighet mellom de to utdanningene for dette utsagnet, nærmere 25%.

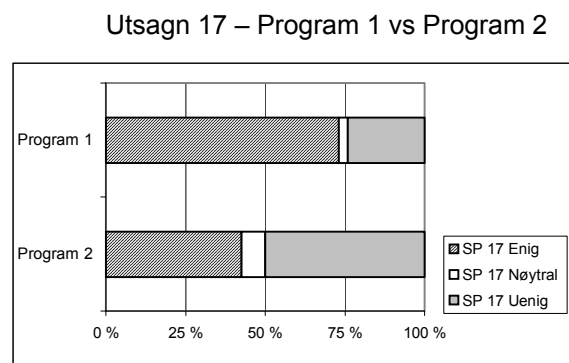
I denne sammenhengen er det også av interesse å se nærmere på *utsagn 17*, som handler om hvordan studentene opplever *kvaliteten* på lærernes feedback; om den er nyttig og til hjelp – i originalversjon: ”Teaching staff here normally give helpful feedback on how you are going”. Et utsagn som vanligvis skiller ”gode” kurs og utdanninger fra de ”dårligste” (Ramsden 1992:107; jf Gibbs & Simpson 2004-5).

Figur 7 viser frekvensmønstrene for utsagn 17; verdien for utsagnet ligger marginalt under GU-skåren for program 1, men ligger noe over for program 2. Imidlertid indikerer samlet resultat for utsagn 7 og 17 at det er betydelig rom for utvikling og forbedring av lærernes feedbackpraksis. Særlig gjelder det utdanningen med lavest prosentvis enighet på begge utsagn, program 2.

Eksemplet representerer en konkretisering av Patricks (1999:35) poeng om at “CEQ results can give us messages we need to hear and which are worth exploring further”. De presenterte resultatene kunne ha ansporet lærerstabene til å analysere mer inngående hvordan de faktisk praktiserer tilbakemelding på studenters arbeid, hvorfor studentene opplevde den aktuelle praksisen som de gjorde – og hvilke former for tilbakemeldinger studentene opplevde som nyttige. I denne forstand kunne CEQ/EMS fungert komplementært til lærer(e)s utviklingsorienterte evalueringsstrategier for å få et bedre grep om dette undervisningsaspektet.



Figur 6



Figur 7

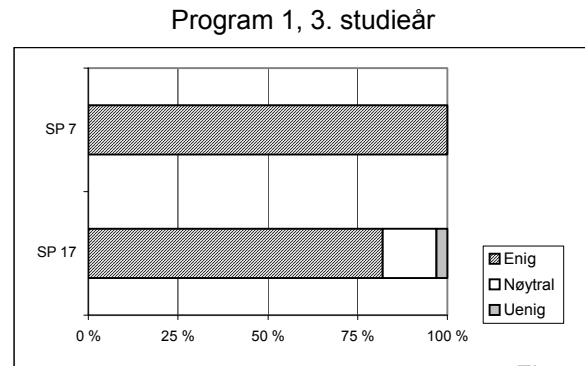
Resultatene fra EMS30 ble ikke anvendt konkret som grunnlag for konsultasjon med de to lærerstabene. Begge studentgrupper ble imidlertid fulgt opp med EMS30 mot slutten av studiet, i vårsemesteret tredje studieår. Resultatene fra denne studien viste en markant positiv endring i studentenes opplevelse av lærerstabens feedbackpraksis i program 1, jf. figur 8.

I løpet av to år økte prosentvis enighet i forhold til mengden feedback (utsagn 7) fra 50 til 100%, jf. figur 7 og 8. For samme periode økte enigheten ift. kvaliteten på lærerstabens feedback fra 73 til 83%. Andelen som opplevde forholdet negativt var redusert fra 24 til 3%. For

studentene i utdanningsprogram 2 viste resultatene motsatt utviklingstendens: Prosentvis enighet for mengden feedback (utsagn 7) var redusert fra 27 til 19%, og for kvaliteten på tilbakemeldingene (utsagn 17) sank prosentvis enighet fra 46 til 13%.

Ut fra analyseeksempelet ser det ut som lærerstaben i program 1 var vesentlig mer oppmerksomme på og lyktes bedre med sin feedback og formative vurdering gjennom hele studieforløpet, enn lærerstaben i program 2. EMS-data som gjelder generell tilfredshet med kvaliteten på studiet (spm 31), viser at prosentvis enighet økte fra 97 til 100% fra første til tredje studieår for program 1. For program 2 sank prosentvis enighet fra 81 til 45%.

Samlet har EMS-resultatene i studien en viss eksemplarisk verdi; de peker i retning av et forhold en av "PBL-pionerene", Howard S. Barrows, løfter fram slik: "Poorly designed, or incompletely designed, problem-based learning curricula can be a frustrating and exhausting experience for both students and faculty alike. Well designed ... problem-based learning curricula can be an exhilarating and rewarding experience for both faculty and students" (Barrows 1985:1).



Figur 8

DEL 3

Validering av *Erfaringer med studiet* (EMS)

Tre studier med norske versjoner av CEQ

Validering av *Erfaringer med studiet*

Utvalgte valideringsstudier med CEQ – en kort oversikt

Det er gjennomført relativt mange valideringsstudier med CEQ i ulike utdanningskontekster, særlig med de to *korte* versjonene, CEQ23 og CEQ24. Her omtales valideringsstudier av den *lange* CEQ-versjonen først, siden de har særlig relevans for validering av både EMS30 og EMS30–GF. CEQ36 inneholder som tidligere nevnt samtlige utsagn fra CEQ30 pluss skalaen *Generelle ferdigheter*. Når det gjelder Ramdens opprinnelige versjon (CEQ30), peker Byrne og Flood (2003) på at det er “... surprisingly little research reporting the replication of the original study ... or the validation of the instrument for use in other contexts” (:128). To vanlige referanser i denne sammenhengen er en studie av Trigwell og Prosser (1991) som omfatter 55 australske sykepleierstudenter, og Richardsons (1994) valideringsstudie basert på 89 studenter. Selv om begge studier i *grove trekk* bekrefter CEQ30s skalastruktur, konkluderer Wilson et al. (1997) med at: ”The confidence with which both these studies can be regarded is limited by the small sizes of their samples”. Valideringsstudien av EMS30 bidrar slik sett til å komplettere bildet; studien bygger på en oversettelse av originalen, er basert på et relativt stort utvalg (n=1.611) og utført i en annen kontekst: Studenter som følger *norske* helse- og sosialfaglige studier, som i hovedsak er helt eller delvis basert på problembaserte prinsipper.

Validering av CEQ36

Wilson et al. (1997) har utført en omfattende valideringsstudie av CEQ, der både CEQ36 og CEQ23 inngår. For CEQ36 bygger valideringen på to større utvalg fra et australsk universitet. Det ene utvalget er på 2.130 studenter fra et bredt utvalg fag- og studieretninger. Det andre utvalget omfatter studenter på bachelornivå (1.–3. år) ved samme universitet; n=7.370 som er 55% av studentpopulasjonen.

Lizzio et al. (2002) har på grunnlag av data fra et av utvalgene over (n=2.130) foretatt mer inngående analyser av sammenhengene mellom studentkarakteristika, opplevd undervisningskvalitet, læringstilnærminger og læringsresultater ved å anvende sti- og regresjonsanalyse. Resultater fra denne studien kommenteres mer detaljert i oppsummeringen av studiene med de tre EMS-versjonene.

John T. E. Richardson har sammen med kollegaer validert CEQ36 i en versjon som er tilpasset fjernstudier/nettbaserte undervisningsprogrammer (Lawless & Richardson 2002; Richardson 2005; Richardson & Price 2003; Richardson & Woodley 2001). I disse studiene varierer størrelsen på utvalgene, fra n=241 til n=2.152. Den opprinnelige skalastrukturen i CEQ-36 er beholdt, men skalaen for god undervisning er delt i to sub-skalaer for å tilpasse spørreskjemaet til nettbaserte studier. De to del-skalaene betegnes henholdsvis "Good tutoring" (5 item) og "Good materials" (3 item), som også empirisk danner to distinkte faktorer (Lawless & Richardsons 2002). Denne CEQ-versjonen representerer med andre ord en variant av CEQ36 med *sju* skalaer.

Validering av CEQ23/24

CEQ23/24 er validert i ulike utdanningssammenhenger med utvalg av varierende størrelse. Wilson et al.s (1997) studie med CEQ23 er basert på responser fra 1.326 avgangsstudenter ved et australsk universitet. Kreber (2003) har validert samme versjon ved et kanadisk universitet med et utvalg studenter på bachelornivå i naturvitenskapelige fag, n=1.080 (undergraduate science students), mens Byrne og Floods (2003) valideringsstudie av CEQ23 baserer seg på et utvalg studenter i merkantile fag (accounting programmes) ved Universitetet i Dublin (n=204).

CEQ24 er også validert i norsk høgskolesammenheng. Espeland og Indrehus (2003) administrerte en norsk oversettelse til et utvalg med 276 sykepleiestudenter fra tre ulike høgskoler, sammen med en norsk oversettelse av *Nursing Clinical Facilitators Questionnaire*. Hensikten var å kart-

legge sykepleierstudentenes vurdering av kvaliteten både i forhold til teoriundervisning og studentenes praksisstudier.

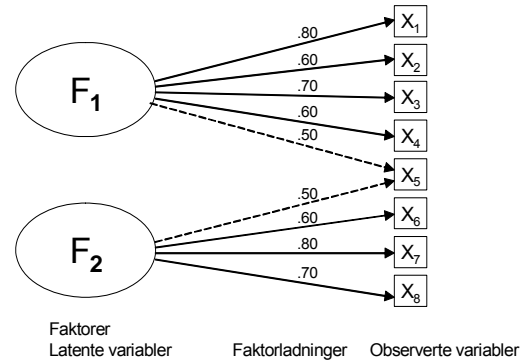
Broomfield og Bligh (1998) har validert CEQ24 med et mindre utvalg medisinerstudenter (n=180), og Steele, West og Simeon (2003) validerte samme versjon ved medisinerutdanningen ved University of the West Indies (n=165). Bligh, Lloyd-Jones og Smith (2000) og Lyon og Hendry (2002) anvendte CEQ24 for å undersøke effekten av omleggingen til problembaserte studiemodeller ved medisinerutdanningene ved henholdsvis University of Liverpool og University of Sydney.

Samtlige studier bekrefter jevnt over CEQ36 og CEQ23/24 som robuste evalueringsverktøy med gjennomgående gode psykometriske egenskaper i ulike fag- og utdanningskontekster. Resultatene for CEQ23/24 er stort sett i overensstemmelse med valideringsdata som årlig rapporteres til GCCA basert på Graduate Destination Survey (se f. eks Ainley 2001).

Kort om faktoranalyse og begrepsvaliditet

For å dokumentere tester og spørreskjemaers begrepsvaliditet anvendes faktoranalyse som statistisk metode. En faktor er en ikke-observerbar, *latent variabel*, utledet av teoretiske begrep (konstrukt); som for eksempel skalaene i EMS: God Undervisning, Klare mål, osv. Det som *er* observer- og målbart, er variablene (utsagnene) i spørreskjemaet, og grunnantakelsen i faktoranalyse er at variasjon i de observerte variablene, reflekterer og bestemmes av variasjonen i en eller flere felles, underliggende faktorer, dvs. i de latente variablene. Validering eller dokumentasjon av begrepsvaliditet dreier seg derfor kort og godt om å påvise at denne antakelsen holder. Metodisk betyr det å undersøke om forhold som antas å henge sammen, faktisk gjør det: At samvariasjon (korrelasjon) mellom antatt, underliggende faktorstruktur og observerbare variabler (spørsmål/ utsagn) har et *akseptabelt nivå*, og at sammenhengene er tilstrekkelig *entydige* (jf. Bryman & Cramer 2005; Hair et al.1998; Kim & Mueller 1978; Loelin 2004; Nordvik & Ulleberg 2000; Pallant 2005; Ulleberg & Nordvik 2001).

Stidiagrammet i figur 9, som viser en to-faktor løsning, illustrerer hovedpoengene i faktoranalyse. Pilene fra faktorene, F_1 og F_2 , betyr at variasjon i faktorene har betydning for variabelskårene. Faktorladninger (egentlig korrelasjonskoeffisienter) viser størrelsen på denne effekten. For utvalg ≥ 250 regnes lad-



Figur 9

ninger større enn 0,30 som *statistisk* signifikante, selv om det ikke er helt entydige regler for forholdet mellom utvalgsstørrelse og grenseverdier for statistisk signifikans. Imidlertid ansees ladninger svakere enn (\pm) 0,50 å ha mindre praktisk signifikans (Hair et al. 1998). Ut fra dette representerer variablene X_1 – X_4 i figur 9 valide operasjonaliserte uttrykk for begrepet som definerer faktoren F_1 .

To piler i figur 9 er stiplet for å løfte fram et forhold betegnet *faktor-kompleksitet*. Variabel X_5 har faktorladninger fra både faktor F_1 og F_2 ; den framstår som tvetydig med en faktorkompleksitet på 2. Ved å utelate variabel X_5 fra spørreskjemaet defineres faktorstrukturen, F_1 og F_2 , entydig ved to sett faktorrene variabler, hhv. X_1 – X_4 og X_6 – X_8 (jf. Ulleberg & Nordvik 2001).

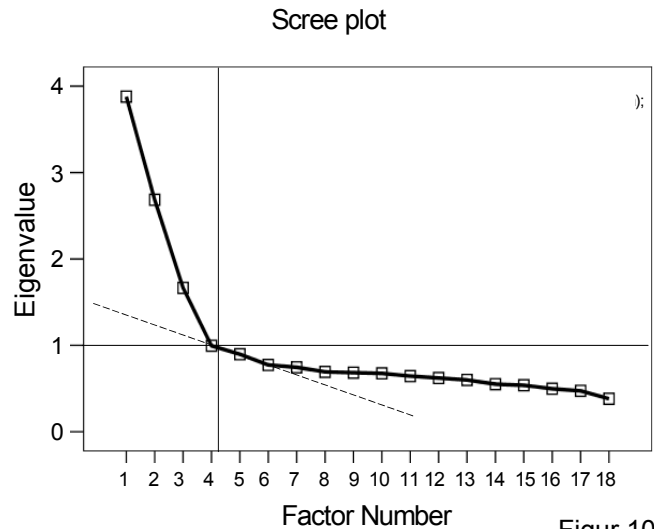
I valideringsstudiene av EMS anvendes *eksplorerende* faktoranalyse, som prinsipielt innebærer at man *ikke* på forhand gjør antakelser om hvor mange underliggende faktorer som trengs for å forklare variasjonen i de observerte variablene på en tilfredsstillende måte. I *konfirmerende* faktoranalyse derimot bestemmes antallet faktorer på forhand, og ved hjelp av mer avanserte metoder som strukturelle ligningsmodeller (structural equations models, SEM) spesifiseres også hvilke variabler som skal ha ladninger fra de ulike faktorene (ibid).⁶

Første trinn i eksplorerende faktoranalyse er å bestemme hvor mange faktorer som trengs for en adekvat reproduksjon av korrelasjonsmatrisen,

⁶ SPSS håndterer ikke denne typen analyser, som krever tilgang til statistikkprogrammer som f. eks LISREL, EQS eller Amos.

som er utgangspunkt for analysen. To vanlige kriterier for faktoruttrekning er *Kaisers kriterium* og visuell inspeksjon av et såkalt *scree plot* ('scree test'). Kaisers kriterium tilsier at faktorer med *eigenvalue* større enn 1 trekkes ut; der *eigenvalue* er et uttrykk for hvor stor andel av variansen i variablene den enkelte faktor forklarer (=summen av kvadrerte faktorladninger).

Scree plot er en grafisk presentasjon av faktorenes *eigenvalue*, som vist i figur 10. "Scree" betyr egentlig "skred", eller "en steinete skråning opp mot en fjellside".⁷ Kriteriet ble introdusert av R. B. Cattell (1966), og ideen er at faktorer som ligger nedenfor et markant brudd i grafen (vist med stiplet linje) forklarer svært lite av variansen i de aktuelle variablene.



Figur 10

I dette eksemplet bør fire faktorer trekkes ut i henhold til Cattells kriterium. Imidlertid har den fjerde faktoren i figur 10 *eigenvalue* < 1 og bør derfor ikke tas med ut fra Kaisers kriterium. Det illustrerer at faktoruttrekning ikke alltid bestemmes entydig av kriteriene: Å bestemme antallet faktorer krever overveielse og tolkning; det handler om å "jakte" på enkle faktorstrukturer som gir en adekvat reproduksjon av opprinnelig korrelasjonsmatrise, på den ene siden – og som samtidig gir en meningsfull fortolkning av de begrepsmessige forhold faktorene referer til, på den andre.

Flere metoder kan anvendes for faktoruttrekning i eksplorerende faktoranalyse, de fire mest kjente er *Principal components analysis* (PCA),

⁷ "... the geological term for the rubble of boulders and debris extended out from the base of a steep mountain slope. The idea is that when you climb up to the top of the scree, you have reached the real mountain slope – the real factors. Below that you have a rubble of trivial or error factors" (Loelin 2004:166). (Etymologisk opprinnelse: scree \approx gammelnordisk, skriða \approx skred).

Maximum likelihood method, Principal axis factoring og *Alpha factoring*. PCA er imidlertid ikke en faktoranalytisk metode i egentlig forstand, men både framgangsmåte og resultater er så pass like for de fire metodene at PCA i praksis regnes som faktoranalyse (jf. Ulleberg & Nordvik 2001:16).

Det neste steget i analysen er *faktorrotasjon*; hensikten med denne operasjonen er å forenkle faktormønsteret slik at faktorene framstår tydeligere og lettere å tolke. To typer akserotering anvendes, hhv. *ortogonal rotasjon* der det forutsettes at faktorene er *ukorrelererte*, og *oblik rotasjon* som tillater at faktorene korrelerer.

I studie A, B og C er SPSS 13.0 anvendt i samtlige analyser.

Studie A

Erfaringer med studiet – versjon EMS30

EMS30 er oversatt fra Ramsdens *Course Experience Questionnaire* slik spørreskjemaet er gjengitt i *Learning to teach in higher education* (1992, i dansk versjon 1999). Denne EMS-versjonen ble administrert til studenter ved tre høyskoler i perioden 2000–2004 som et ledd i halvårige/årlige studentevalueringer av undervisning og læringsmiljø. Respondentene kommer fra 35 ulike klasser/kull i helse- og sosialfaglige utdanninger (barnevern-pedagoger, bioingeniører, sosionomer, sykepleiere og videreutdanningsstudenter i helse- og sosialfag). Studien er basert på 1.611 komplett utfylte EMS-spørreskjemaer. Respondentene fordeler seg slik: 38% studenter fra 1. studieår, 15% fra 2. studieår, 34% fra 3. studieår og 13% fra 4. studieår (videreutdanninger). De tre høyskolene er representert slik: 73 % studenter fra *Høgskole I*, 14% fra *Høgskole II* og 13% fra *Høgskole III*. Kjønnssfordeling i utvalget er 72% kvinner og 28% menn, og gjennomsnittsalder er 25,7 år.

Begrepsvaliditet

Mønstermatrise fra *eksplorerende* faktoranalyse basert på uttrekning med *prinsipale komponenters metode* (PCA) og *oblik* akserotasjon er gjengitt i tabell 4; ladninger mindre enn 0,30 er utelatt i tabellen. Løsningen med fem uttrukne faktorer er basert på inspeksjon av scree plot; Kaisers kriterium tilsier at seks faktorer bør trekkes ut (eigenvalue >1). En femfaktor-løsning stemmer imidlertid best med EMS' underliggende begrepmessige faktorstruktur, og de fem faktorene forklarer til sammen 44,5% av total varians. I matrisen presenteres variablene (utsagnene) i fallende orden etter størrelsen på faktorladninger fra antatt underliggende faktor; ladninger i samsvar med antatt faktorstruktur er kursivert. Korte tekster fra hvert item indikerer forholdet som den enkelte variabel måler; tallene i parentes viser det enkelte items plassering i spørreskjemaet (jf. vedlegg 1).

Tabell 4

Faktoranalyse, mønstermatrise for EMS30, n=1611

Antatt faktorstruktur	Variabler/item	FAKTORANALYSE				
		Uttrukne faktorer				
		1	2	3	4	5
GU (27): lærerne forsøker å få alle til å yte sitt beste	0,71					
GU (3): motiverer meg til å gjøre mitt beste	0,64					
GU (7): bruker tid på å kommentere/tilbakemeldinger	0,62					
GU (15): legger vekt på å forstå vansker og problemer	0,64					
GU (18): lærerne er virkelig gode til å forklare	0,62					
GU (17): tilbakemeldinger om hvordan jeg ligger an	0,58					
GU (20): jobber for gjøre lærestoffet interessant	0,58					
GU (25): lærerstaben viser [ingen] virkelig interesse.....*)	0,45		0,41			
KM (6): klar forestilling om hvordan jeg greier meg					-0,75	
KM (1): lett å forstå og finne ut av krav som stilles					-0,68	
KM (13): [ikke] vanskelig å finne ut hva som forventes	*)				-0,63	
KM (29): lærerne gjør det helt klar fra begynnelsen	0,41				-0,44	
KM (19): læremålene er [ikke] klart formulert	*)			-0,32	-0,36	
ARB (22): [ikke] stort press/høye krav	*)	0,77				
ARB (4): arbeidsbyrden er for stor	*)	0,72				
ARB (30): omfanget gjør at vi [ikke] kan lære alt grundig	*)	0,55				
ARB (14): vi får tid nok til forstå tingene		0,52				
ARB (10): undervisningsplanen dekker [ikke] for mye	*)	0,31	-0,44			
VF (12): lærerne [ikke] mer interessert i å teste	*)		-0,61			
VF (8): det holder [ikke] med en god hukommelse	*)		-0,59			
VF (5): lærerne mener de [ikke] kan lære noe av oss	*)		-0,52			
VF (21): lærere stiller bare spørsmål om faktakunnskaper	*)		-0,50			
VF (26): kommer seg [ikke] gjennom bare ved å pugge	*)	-0,39	-0,47			
VF (23): tilbakemeldinger består [ikke] bare av rettinger	*)	0,39				
SA (16): stor valgfrihet i forhold til oppgaver og studiearbeid					0,66	
SA (11): muligheter til å styre eget studiearbeid og læring					0,64	
SA (2): [få] muligheter til å fordype seg ut fra interesse	*)				0,60	
SA (28): det er [få] valgmuligheter mht. vurderingsformer	*)				0,31	
SA (24): diskuterer med lærerne hvordan vi bør lære	0,64					
SA (9): blir oppmuntret til å utvikle egne faglige interesser	0,57					

*) Viser til negativt formulerte utsagn som er invertert i analysen; i tabellen er dette markert språklig med 'ikke' eller 'få' i hakeparentes, unntatt for utsagn ARB4 og VF21. Halvparten av utsagnene i EMS30 er negativt formulert.

Mønstermatrisen i tabell 4 bekrefter i all hovedsak EMS' underliggende faktor- og begrepsstruktur; i alt 27 variabler får signifikante ladninger fra faktorer i samsvar med EMS' faktorstruktur. Det er *tre* variabler som bryter med strukturen: Utsagnene VF23, SA9 og SA24 får entydige, signifi-

kante faktorladninger fra faktor 1 (God undervisning). I tillegg har *fem* variabler doble faktorladninger (faktorkompleksitet på 2); for utsagnene GU25 og KM29 er ladninger fra antatt og alternativ faktor av omtrent samme styrke. Imidlertid har de faktorkomplekse variablene ladninger av mindre praktisk betydning; samtlige ligger i området 0,32–0,44 (jf. Hair et al. op.cit). Det er med andre ord skalaen *Studentautonomi* som avviker tydeligst fra EMS' underliggende faktorstruktur med to variabler (utsagn 9 og 24) som får faktorrene, signifikante ladninger fra *God undervisning*.

Mer detaljerte kommentarer til de enkelte faktorene:

- *Faktor 1* forklarer 22,3% av total varians; den identifiseres som *God undervisning* med signifikante ladninger til *samtlige* utsagn. Faktor 1 har i alt signifikante ladninger til 12 variabler, en variabel fra hhv. Klare mål (KM29) og Vurderingsformer (VF23) og to fra Studentautonomi (SA 9 og 24) – i tillegg til utsagnene i GU. I alt får *sju* av åtte utsagn i God undervisning signifikante og faktorrene ladninger fra faktor 1.
- *Faktor 2* forklarer 7,6% av samlet varians; den identifiseres som *Arbeidsbyrde*: Alle utsagn i skalaen får signifikante ladninger fra faktor 2. Et av utsagnene har faktorkompleksitet på 2. Oppsummert er det *fire* av fem utsagn i skalaen Klare mål som har entydige og signifikante ladninger fra faktor 2; i tillegg har faktor 2 en signifikant ladning til Vurderingsformer.
- *Faktor 3* forklarer 5,8 % av samlet varians; den identifiseres som *Vurderingsformer* med signifikante ladninger til *fem* av skalaens seks utsagn. Faktor 3 har i alt signifikante ladninger til åtte variabler; i tillegg til utsagnene i VF gjelder det en variabel fra hhv. God undervisning, Klare mål og Arbeidsbyrde. Det er *fire* av seks utsagn i Vurderingsformer som får entydige, signifikante faktorladninger fra faktor 3. Ett av de øvrige to utsagnene har dobbel faktorladning, det andre får entydig, signifikant ladning fra God undervisning.

- *Faktor 4* forklarer 4,9 % av samlet varians; den defineres som *Klare mål*: alle fem utsagnene får signifikante ladninger fra faktor 4, hvor tre av disse får signifikante og entydige ladninger denne faktoren. De to øvrige utsagnene har faktorkompleksitet på 2.
- *Faktor 5* forklarer 3,8% av total varians; den identifiseres som *Studentautonomi*: Fire av skalaens seks utsagn får entydige, signifikante ladninger fra faktor 5. De øvrige to utsagnene har entydige og signifikante ladninger fra faktor 1, God undervisning.

Fullstendig og entydig dokumentasjon av antatt, latent faktorstruktur er snarere unntaket enn reglen i faktoranalytiske valideringsstudier. Avvikene i tabell 4, i form av kryssladninger og doble faktorladninger, har likhetstrekk med resultater dokumentert i flere større valideringsstudier med den engelske versjonen, CEQ, blant annet:

- (a) I Ramdsens (1991) pilotstudie mottar *to* variabler fra God undervisning større ladninger fra Vurderingsformer enn fra GU; det gjelder også en av variablene i Studentautonomi.
- (b) I Lawless og Richardsons (2002) studie med CEQ36 viser de ”problematiske” variablene KM29, VF23 og SA24 samme mønster som i analysen av EMS30 ved at de får signifikante ladninger fra God undervisning.
- (c) I Wilson et al.s (1997) studie med CEQ36 er det variablene VF23, VF26, SA9 og SA24 som viser et tilsvarende mønster: enten ved kryssladninger eller faktorkompleksitet på 2.

Flere av utsagnene nevnt over, sammen med hele skalaen Studentautonomi, er utelatt fra CEQ23, for å styrke spørreskjemaets psykometriske egenskaper. Poenget er illustrert i tabell 5, som viser mønstermatrise basert på faktorisering av de 17 utsagnene i EMS som er videreført til CEQ23. Analysen er basert på prinsipale komponenters metode med oblik rotasjon, og de fire uttrukne faktorer forklarer 52,1% av samlet varians. Faktorladninger mindre enn 0,30 er utelatt i tabellen.

Tabell 5 Mønstermatrise for 17 utsagn fra EMS30 videreført i CEQ23 (n=1611)

Antatt faktorstruktur	Uttrukne faktorer			
	1	2	3	4
GU7	0,72			
GU15	0,69			
GU18	0,68			
GU17	0,66			
GU3	0,65			
GU20	0,62			
KM6			0,78	
KM1			0,71	
KM13			0,69	
KM29	0,37		0,44	
ARB22		0,80		
ARB4		0,76		
ARB30		0,59		
ARB14		0,57		
VF12				0,68
VF8				0,66
VF5				0,63

Resultatet i tabell 5 viser en tilnærmet faktorren løsning; unntaket er utsagn KM29 som har faktorkompleksitet 2 med ladning fra faktor 1, identifisert som God undervisning.

Reliabilitet – skalaenes indre konsistens

Reliabilitetskoeffisienter for skalaene i EMS30 uttrykt ved *Cronbachs alfa* (indre konsistens) er oppsummert i tabell 6. Kolonnen til høyre i tabellen viser koeffisienter for tilsvarende skalaer i to studier med CEQ36, fra Wilson et al. (op cit).

Tabell 6 Cronbachs alfa for skalaene i EMS30

Skalaer	EMS30	CEQ36	
	2000–04	1993	1994
God undervisning (GU)	0,85	0,86	0,88
Klare mål (KM)	0,65	0,82	0,82
Arbeidsbyrde (ARB)	0,65	0,75	0,74
Vurderingsformer (VF)	0,60	0,74	0,73
Studentautonomi (SA)	0,58	0,68	0,67

Tabellen viser høy reliabilitet for *God undervisning* i samtlige studier, $\alpha=0,85\text{--}0,88$. De øvrige skalaene i EMS har fra moderate til relativt lave alfa-verdier (fra 0,65 til 0,58), gjennomgående lavere enn resultatene fra de australske undersøkelsene (fra 0,82 til 0,67).

Det er imidlertid et spørsmål hvor høye alfakoeffisienter som bør kreves av de ulike skalaene i EMS; hva som bør regnes som grense for akseptabel reliabilitet. Litteraturen opererer med ulike generelle ”tommelfingerregler”, i følge Hair et al. (1998:118) er: ”The general agreed upon lower limit for Cronbach’s alpha ... [is] .70, although it may decrease to .60 in exploratory research.” Pedhazur og Schmelkin (1991:101f) problematiserer verdien av slike autoritative standardanvisninger uten nærmere overveielser og vurderinger.

Antall utsagn er et forhold som påvirker skalaers indre konsistens: Et velprøvd råd for å øke Cronbachs alfa, særlig for korte skalaer, er å øke antallet item. Gevinstene er imidlertid marginale ved økning utover 10–12 item. Legges gjennomsnittlig inter–item-korrelasjon til grunn for reliabilitetsvurderinger, som prinsipielt er uavhengig skalalengde, anses korrelasjonskoeffisienter rundt 0,30 som fullt ut akseptabelt. En gjennomsnittlig inter–item-korrelasjon på 0,30 gir følgende alfaverdier for skalaer med henholdsvis 5, 6 og 8 item: 0,68 – 0,72 – 0,77. I forhold til dette har skalaene *God undervisning*, *Klare mål* og *Arbeidsbyrde* akseptabel reliabilitet, mens skalaene *Vurderingsformer* og *Studentautonomi* nærmer seg grenseområdet for det som vanligvis regnes som akseptabel reliabilitet, når en mer rigorøs og stringent tolkning av psykometriske standarder legges til grunn, (jf. Christophersen 2004; Hair et al. 1998; Pedhazur & Schmelkin 1991; Smith 1981; se også vedlegg 4: Oversiktstabell for sammenhenger mellom inter–item–korrelasjon, Cronbachs alfa og skalalengde/antall item, basert på Spearman–Brown-formelen).

Studie B

Erfaringer med studiet – versjon EMS30–GF

I EMS30–GF erstatter skalaen *Generelle ferdigheter* (GF) skalaen *Student-autonomi*, derfor betegnelsen EMS30–GF; de øvrige utsagnene er de samme som i EMS30. Versjonen ble prøvd ut som et ledd i utviklingen av høgskolens kvalitetssystem. Ideen var å teste ut denne versjonen som en mulig felles indikator for studentenes vurdering av undervisningskvalitet og læringsmiljø på høgskolens ulike utdanningsprogrammer. Intensjonen var også å teste hvordan skalaen *Generelle ferdigheter* fungerte psyko-metrisk i denne høgskolekonteksten.

EMS30–GF ble administrert til studenter ved tre avdelinger ved høgskolen i vårsemesteret 2005: (1) ingeniørutdanningen; (2) lærerutdanningen; (3) utdanninger i språk, samfunnsfag, samt organisasjon og ledelse. Til sammen 317 komplett utfylte spørreskjemaer ble returnert. Respondentene for-deler seg slik på de ulike studieprogrammene:

- | | |
|----------------------------------|---|
| • Ingeniørutdanning: | 60 studenter fordelt på 1., 2. og 3. studieår |
| • Førskolelærerutdanning: | 47 studenter fordelt på 1. og 2. studieår |
| • Lærerutdanning: | 47 studenter fordelt på 1– 4. studieår |
| • Praktisk–pedagogisk utdanning: | 40 studenter |
| • Bachelorprogram, språk: | 50 studenter i 1.år, engelsk og fransk |
| • Bachelorprogram, økonomi | 56 studenter i 2. studieår |
| • Masterprogram, organisasjon: | 17 studenter |

De enkelte utdanningene foresto studentutvalg og administrasjon av EMS30–GF, der man la vekt på å trekke ut et rimelig representativt utvalg fra de enkelte utdanningsprogrammene. Kjønnssfordeling i hele utvalget: 58% kvinner og 42% menn. Gjennomsnittsalder er 28,1 år for de respon-dentene som oppga alder (n=304).

Begrepsvaliditet

Tabell 7 viser mønstermatrisen fra *eksplorerende* faktoranalyse; faktoruttrekning med *prinsipale komponenters metode* og *oblik rotasjon*. Faktorladninger mindre enn 0,30 er utelatt (for øvrig er tabellen satt opp som tabell 4 og 5). Kaisers kriterium viser at sju faktorer skal trekkes ut; imidlertid er det en tendens til ”overfaktorerings” når dette kriteriet følges fullt ut. Scree plot er noe tvetydig, men antyder uttrekning av 4–6 faktorer. Både en fire- og seksfaktorløsning produserer imidlertid lite tolkbare matriser i forhold til EMS–30GFs underliggende struktur. Løsningen med *fem* faktorer vist i tabell 7 gir beste løsning for det aktuelle utvalget.

Tabell 7

Skala- struktur	Mønstermatrise for EMS30–GF, n=317				
	FAKTORANALYSE				
	Uttrukne faktorer				
	1	2	3	4	5
GU (17)	0,72				
GU (15)	0,69				
GU (27)	0,67				
GU (7)	0,66				
GU (20)	0,66				
GU (18)	0,65				
GU (3)	0,64				
GU (25)	0,58				
GF (16)			0,76		
GF (28)			0,69		
GF (24)			0,66		
GF (2)			0,62		
GF (9)			0,52		
GF (11)			0,38		
KM (6)				0,79	
KM (1)				0,74	
KM (13)				0,63	
KM (29)	0,35			0,53	
KM (19)	0,38				
ARB (4)		0,74			
ARB (22)		0,65			
ARB (30)		0,58			
ARB (14)		0,56			
ARB (10)		0,51			
VF (8)					0,80
VF (12)					0,66
VF (21)					0,59
VF (26)		-0,38			0,35
VF (5)	0,51				
VF (23)			-0,36		

Mønstret som avtegnes i tabell 7 er i hovedsak sammenfallende med analysen av EMS30 for skalaene som inngår i begge EMS-versjoner. Det gjelder særlig de 17 utsagnene som er videreført i CEQ23, jf. tabell 5 (s. 50).

Analysen viser at samtlige utsagn mottar signifikante ladninger fra de fem faktorene, som til sammen forklarer 48% av total varians – fordelt slik: 22% forklares av faktor 1; faktor 2 forklarer 8,5%; faktor 3 forklarer 6,9 %; faktor 4 forklarer 6,1% og faktor 5 forklarer 4,8 %.

Faktor 3 identifiserer *Generell ferdigheter* hvor alle item får signifikante, entydige faktorladninger fra denne faktoren. Faktor 3 har i tillegg en mindre, (negativ) signifikant ladning til et utsagn på skalaen Vurderingsformer, VF23, som handler om formativ feedback, og gjelder hvorvidt tilbakemeldinger på studentenes arbeider utelukkende gis som karakterer og/eller korte, skriftlige kommentarer. I det gjeldende studentutvalget oppfattes fraværet av mer utfyllende og detaljert formativ tilbakemelding som noe hemmende for utvikling av generelle ferdigheter og kompetanseinnretninger.

Skalaen *Generelle ferdigheter* har en nærmest entydig, faktorren løsning (basert på signifikante faktorladninger). Sammenholdt med resultatene i studie A gir det grunn til å anta at en norsk versjon av CEQ23 også vil vise tilfredsstillende begrepsvaliditet i norske høyskolekontekster (slik Espeland og Indrehus' (2003) resultater tyder på; jf. studie C).

Reliabilitet – indre konsistens

Tabell 8 viser Cronbachs alfa for skalene i CEQ30–GF; reliabilitetskoeffisienter for EMS30 er gjengitt i høyre kolonne for sammenlikning.

Tabell 8

Cronbachs alfa for EMS30–GF

Skalaer	EMS30–GF	EMS30
GU: God undervisning	0,85	0,85
GF: Generelle ferdigheter	0,74	--
KM: Klare mål	0,72	0,65
ARB: Passende arbeidsbyrde	0,67	0,65
VF: Vurderingsformer	0,52	0,60

God undervisning har høy reliabilitet, som i studie A. Også *Generelle ferdigheter* viser god indre konsistens, og for *Klare mål* er alfaverdien vesentlig høyere enn for EMS30 i studie A. Endringen for *Arbeidsbyrde* er marginal, mens alfaverdien for *Vurderingsformer* er påfallende lav; det er primært utsagn VF23 som bidrar til dette. Alfaverdien for *Vurderingsformer* er 0,63 når dette itemet ikke regnes med.

En merknad om *Generell ferdigheter* og CEQ23

Revisjonen av den originale CEQ30 og utviklingen av CEQ23 medfører et par problematiske poenger som bør løftes fram. Det første gjelder skalaen *Generelle ferdigheter*s begrepsmessige status i forhold til de øvrige skalaene. Denne skalaen er som Sadlo og Richardson (2003) påpeker, "... rather different from the original scales, in that it requires students *to evaluate the skills that they personally have gained* from their programmes, rather than *the quality of the programmes themselves*" (:256; min kursivering). Den fokuserer mer på studenters *vurdering* av antatte læringsresultater eller læringsprodukter enn på deres *opplevelse* og *erfaring* med undervisnings- og læringsprosessen.

Det andre forholdet knytter seg til et mer allment dilemma i evalueringsforskning: Spørsmålet om hvorvidt evalueringsresultatene bør presenteres som "... a single score, a set of factor scores, ... or responses to individual items (Marsh 1987:260).

Kritikken som rettes mot faktorbaserte spørreskjemaer generelt og CEQ spesielt, er at de "... provide only a limited coverage of the student experience" (Wilson et al. 1997:35). Implisitt framheves evalueringsverktøy basert på flere og mer spesifiserte enkelt-item som alternativ, fordi dette gir mer detaljert informasjon og et bredere erfaringsgrunnlag, særlig med sikte på utvikling av pedagogisk praksis. I denne forstand argumenteres det for å styrke evalueringens *didaktiske relevans*.

Utviklingen av CEQ23 hviler derimot på et psykometrisk "stringensargument", for eksempel formulert slik: "People will only have confidence in CEQ data to the extent that they have confidence in the instrument it-

self” (ibid) – som refererer til de psykometriske overveielser som har ført til at i alt 13 utsagn fra CEQ30 er sløyfet. I et relevansbasert perspektiv bidrar det til å redusere erfaringsgrunnlaget for evalueringen; det gjelder blant annet skalaen *Studentautonomi*. Når den erstattes fullt ut av *Generelle ferdigheter*, utelukkes samtidig data som gjelder studentenes oppfatninger om valgfrihet og relativ autonomi i undervisningssituasjoner og læringsmiljøer. I problem- og praksisbaserte studieprogrammer hvor prinsippet om selvstyrt læring står sentralt, vil imidlertid slike data være av særlig interesse (jf. Pettersen 2005).

Studie C

Erfaringer med studiet – versjon EMS26

EMS26 er oversatt fra CEQ23 og utvidet med tre utsagn fra skalaen Studentautonomi. EMS26 representerer på denne måten et forsiktig kompromiss mellom ”stringens” og ”relevans”, med en intensjon om å styrke CEQ23s didaktiske relevans ved å inkludere alle seks skalaer.

Valideringsstudien bygger på to utvalg. Det ene omfatter grunnutdanningsstudenter ved høgskolens avdeling for helse- og sosialfag, med respondenter fra disse studieretningene: Barnevernspedagoger, bioingeniører, sosionomer, sykepleiere og vernepleiere. Det andre utvalget består av studenter fra ulike studieprogrammer ved fire av høgskolens avdelinger. De besvarte EMS26 som ledd i høgskolens program for avdelingsovergripende studentevaluering av undervisning og læringsmiljø. Totalt 876 komplett utfylte spørreskjemaer inngår i valideringen; tabell 9 viser hvordan respondentene er fordelt i grove trekk. For utvalg 2 er det omtrent like mange studenter fra 1., 2., og 3. studieår.

Tabell 9

Samlet utvalg i studie C, n=876

Utvalg 1 (vårsemester 2005)			Utvalg 2 (vårsemester 2006)		
Avdeling for helse- og sosialfag, n=436			Høgskolens avdelinger, n= 440		
Fordeling på studieår	N		Fordeling på studieprogram	N	
Studenter 1. år	96		Helse- og sosialfag	188	
Studenter 2. år	187		Ingeniør og ITK	148	
Studenter 3. år	153		Lærerutdanning	67	
			Økonomi- og språkfag	37	
	Kvinner	81%		Kvinner	60%
	Menn	19%		Menn	40%
	Gjennomsnittsalder	28 år		Gjennomsnittsalder	29 år

Mønstermatrise fra *eksplorerende* faktoranalyse med *prinsipale komponenters metode* og *oblik rotasjon* er gjengitt i tabell 10. Faktorladninger mindre enn 0,30 er utelatt. Seks faktorer er trukket ut, i overensstemmelse med *både* Kaisers kriterium og inspeksjon av scree plot. Den andre kolonnen i tabell 10 viser utsagnenes plassering i EMS30–GF; SA–utsagnenes plassering i EMS30 står i parentes.

Tabell 10

Faktoranalyse, mønstermatrise for EMS26, n=876

Latent faktorstruktur	Itemnr. i EMS30-GF (i EMS30)	Uttrukne faktorer					
		1	2	3	4	5	6
GU (18)	18	0,77					
GU (15)	15	0,68					
GU (19)	20	0,67					
GU (7)	7	0,65					
GU (3)	3	0,62					
GU (17)	17	0,59					
GF (16)	16				0,71		
GF (22)	24				0,67		
GF (2)	2				0,66		
GF (11)	11				0,66		
GF (9)	9				0,63		
GF (24)	28				0,63		
KM (6)	6			-0,73			
KM (13)	13			-0,72			
KM (1)	1			-0,68			
KM (25)	29	0,31		-0,53			
ARB (21)	22		0,76				
ARB (4)	4		0,71				
ARB (26)	30		0,64				
ARB (14)	14		0,61				
VF (12)	12					0,72	
VF (20)	21					0,62	
VF (8)	8					0,60	
SA (10)	(11)						0,73
SA (23)	(16)						0,70
SA (5)	(2)						0,51

Begrepsvaliditet og reliabilitet

Mønstermatrisen i tabell 10 innfrir forventningene fra studie A og B: Samtlige variabler med ett unntak får entydige, signifikante ladninger fra faktorer som identifiserer EMS26s underliggende teoretiske faktorstruktur. Som i studie A og B er det utsagn 25 i skalaen Klare mål som har faktor-kompleksitet på 2, med en relativt beskjeden ladning (0,31) fra faktor 1 (God undervisning). Dette utsagnet (i originalversjon): ”The staff made it clear from the start what they expected from students” – forbindes med god undervisning i flere studier (se f eks Ainley 2001; Lawless & Richardson

2002; Wilson et al. 1997). I EMS26 er det dette utsagnets tvetydige ladning som teknisk sett hindrer ”full match” med underliggende begrepsstruktur.

De seks skalaenes indre konsistens, Cronbachs alfa, er satt opp i tabell 11 med koeffisienter fra Espeland og Indrehus’ (2003) studie med en norsk versjon av CEQ24,⁸ og alfaverdier fra *Graduate Destination Survey* på bachelornivå (Ainley 2001) som referanser.

Tabell 11

Skalaer	Cronbachs alfa for EMS26		
	EMS26 Studie C	Espeland & Indrehus (2003)	Graduate Survey Ainley (2001)
God undervisning (6 item)	0,82	0,75	0,87
Generelle ferdigheter (6 item)	0,71	0,79	0,77
Klare Mål (4 item)	0,66	0,68	0,78
Passende arbeidsbyrde (4 item)	0,67	0,69	0,71
Vurderingsformer (3 item)	0,56	0,37	0,72
Studentautonomi (3 item)	0,45	--	---
N=	876	276	ca 50.200

Faktorene (skalaene) i EMS26, med unntak av *Studentautonomi*, har gjennomgående akseptabel og tilfredsstillende reliabilitet, fra moderate til høye alfaverdier. Studentautonomi er den svakeste skalaen med $\alpha=0,45$, som tilsvarer en gjennomsnittlig inter–item-korrelasjon på 0,21.

EMS – en generell indikator for opplevd undervisningskvalitet?

En problemstilling som gjerne tematiseres i valideringsstudier av CEQ, dreier seg om hvorvidt de fem, ev. seks, skalaskårene kan summeres og leses som en enhetlig indikator på ”opplevd undervisningskvalitet”, dvs. at summen uttrykker en mer generell, bakenforliggende eller overordnet faktor.

⁸ Når reliabilitet for Vurderingsformer er lav i denne studien, $\alpha=0,37$, kan det skyldes at det fjerde utsagnet i skalaen er tatt med i analysen, jf. Del 1, s. 24. Dette utsagnet får som regel ladning fra God undervisning; i Espeland & Indrehus’ studie har utsagnet signifikant ladning fra Generelle ferdigheter (0,44).

I en tidligere studie hevder Richardson (1994), basert på andre-ordens faktoranalyse, at faktorene i CEQ30 danner et ”monarcial hierarchy”: En generell, overordnet faktor basert på fem korrelerte, men ikke-overlap-pende førsteordens faktorer. Også i Richardsons (2005) validering av den tilpassede versjonen av CEQ36 danner faktorene én høyere ordens faktor; betegnet ”Perceived Academic Quality”, som uttrykk for et felles mål på opplevd undervisningskvalitet.

Det er imidlertid sterke argumenter for at skalaene i CEQ reflekterer *to* høyere ordens faktorer, henholdsvis en som inkluderer God undervisning, Generelle ferdigheter, Klare mål, Vurderingsformer og ev. Student-autonomi, og en faktor som består av Arbeidsbyrde alene. Denne løsningen får sterkest støtte basert på sti- analyser av datamaterialet fra både CEQ36 og CEQ23 (Wilson et al. 1997:42f).

Faktoranalyse basert på faktorskårene i hhv. EMS30, EMS30–GF og EMS26 er oppsummert i tabell 12; Richardsons en-faktorløsning er tatt med som referanse under studie A.

Tabell 12 Faktoranalyse basert på faktorskårer i EMS

	Studie A (EMS30)			Studie B (EMS30–GF)		Studie C (EMS26)	
	EMS30	Richardson CEQ36					
	Faktor 1	Faktor 1		Faktor 1	Faktor 2	Faktor 1	Faktor 2
GU	0,85	0,74/0,58		0,85		0,64	
GF	--	0,60		0,53		0,65	
KM	0,58	0,73		0,54		0,47	
ARB	0,38	0,47			0,74		0,79
VF	0,58	0,53		0,55		0,41	
SA	0,62	0,52		---		0,34	

For studie A, B og C gjelder: Faktortuttrekning, *Principal axis factoring* med oblik rotasjon.
Faktorladninger < 0,30 er utelatt i tabellen.

Analysen av EMS30 støtter Richardsons en-faktorløsning; i tabellen gjelder de to verdiene for GU i Richardsons studie delskalaene ’Good materials’ og ’Good tutoring’. Resultater fra både EMS30–GF og EMS26 gir imidlertid tydelige *to-faktorløsninger*, der Arbeidsbyrde får signifikante ladninger fra faktor 2 på hhv. 0,74 og 0,79.

I en nyere studie med en norsk, forkortet variant av CEQ (Diseth et al. 2006) støttes også primært én høyere ordens en-faktorløsning basert på

konfirmerende faktoranalyse. I denne versjonen er bare 18 utsagn fra CEQ24 videreført, og to nye utsagn tilpasset den aktuelle undervisningskonteksten er lagt til. De 20 utsagnene er fordelt på fire skalaer: GU (6 utsagn), KM (4 utsagn), ARB (6 utsagn) og VF (4 utsagn). Det kan imidlertid tenkes at en såpass omfattende revisjon av CEQs skalastruktur i seg selv er bedre tilpasset en løsning med én høyere ordens faktor. Eksplorative faktoranalyse basert på samme fire delskalaer i EMSGF-30 og EMS-26 underbygger denne muligheten: Begge analyser resulterer i klare en-faktorløsninger.⁹

Curtis og Keeves' (2000) har nærmet seg "single score"-problematikken fra et noe annet perspektiv. Basert på 'Rasch målingsmodell', en relativt avansert statistisk metode, har de undersøkt hvilke CEQ-utsagn som kvalifiserer som variabler i en enhetlig, institusjonell *kvalitetsindikator* (institutional performance indicator), kalt *Graduate Satisfaction Index* (GSI). Basert på analyse av data fra Graduate Destination Survey (n=51.631) tilfredsstiller 17 av itemene i CEQ24 disse kravene. Curtis og Keeves' perspektiv er *kvalitetssikring*; de har primært beslutningstakere i tankene: "If significant policy decisions such as resource allocations are to be based upon instruments such as CEQ, we suggest that better analytic techniques ... should be employed so that those policy decisions are more soundly based" (:82).

⁹ Dette berører også et mer grunnleggende problem knyttet til avkorting av lengre spørreskjemaer: "In deciding how to interpret items in questionnaires ... , participants will make use of the immediate context: in other words, the content of neighbouring items. If the context is changed, then their responses *may well* be changed" (Richardson 2000:110; min kursivering). Derfor bør vi ikke umiddelbart anta at studenter responderer på samme måte på en 20-items versjon av CEQ som de gjør på samme utsagn integrert i f. eks. CEQ36.

EMS – kriterievaliditet

Vanligvis inngår også vurdering av spørreskjemaets *kriterievaliditet* i valideringsstudier av CEQ. Kriterievaliditet gjelder korrelasjon mellom instrumentet/skalaer og andre mål for de aktuelle fenomenene som instrumentet søker å studere/kartlegge. For CEQ handler kriterievaliditet om skalaenes relasjon/korrelasjon med relevante eksterne kriterier og mål knyttet til *undervisning og læring*, hvor følgende kriterier vanligvis inngår (jf. Wilson et al. 1997):

1. Korrelasjon mellom CEQ-skalaene og *studentenes tilfredshet med studiet*.
2. Korrelasjon mellom CEQ-skalaene og studentenes sjølrapportering om bruk av hhv. *dybde- og overflatetilnærming* til studiearbeid og læring.
3. Korrelasjon mellom *Generelle ferdigheter*, som *læringsresultat*, og de øvrige CEQ-skalaene.
4. Korrelasjon mellom CEQ-skalaer og læringsresultater, f. eks. *eksamensresultater*.

Datagrunnlaget i studie A, B og C gir muligheter for analyser av CEQs kriterievaliditet i forhold til pkt 1–3; når det gjelder pkt 4, foreligger ikke eksamens- eller vurderingsdata i de tre studiene.

EMS som uttrykk for studenters tilfredshet med studiet

Den underliggende antakelsen er at CEQs skala- eller faktorskårer, som sider ved opplevd undervisningskvalitet, vil samvarierte med studentenes generelle tilfredshet med kvaliteten på studie- og undervisningsprogrammer. Tilleggsspørsmålet i de ulike variantene av EMS/CEQ danner ytre valideringskriterium, hvor studentenes *tilfredshet/fornøydhet* med det aktuelle undervisningsprogrammet etterspørres – formulert i de fleste CEQ-versjoner som: *In general, I'm satisfied with the quality of this course.*

Tabell 13 viser korrelasjonskoeffisienter for studiene A, B og C. Resultater fra to tidligere omtalte valideringsstudier, Ramsden (1991) og Wilson et al. (1997), er satt opp som sammenlikningsgrunnlag; samtlige koeffisienter i disse studiene er signifikante for $p < 0.001$. Laveste og høyeste verdier oppnådd i de tre undersøkelsene i Wilson et al.s studie er ført opp i kolonnen helt til høyre.

Tabell 13 Korrelasjon mellom skalaene i EMS/CEQ og tilfredshet med lærings- og undervisningsmiljø

EMS-/CEQ-skalaer	Studie A	Studie B	Studie C	Ramsden (1991)	Wilson et al. (1997) 3 studier
	EMS30	EMS-GF	EMS26	CEQ30	CEQ23/36
God undervisning	0,52**	0,62**	0,60**	0,60	0,63–0,64
Klare mål	0,34**	0,51**	0,39**	0,47	0,50–0,55
Arbeidsbyrde	0,20**	0,24**	0,21**	0,21	0,17–0,36
Vurderingsformer	0,37**	0,32**	0,33**	0,40	0,33–0,47
Studentautonomi	0,30**	--	0,28**	0,40	0,51–0,54
Generelle ferdigheter	--	0,37**	0,45**	--	0,47–0,51

** signifikant for $p < 0.01$

Resultatene fra studiene med de tre EMS-versjonene viser et korrelasjonsmønster som jevnt over faller innenfor, eller ligger tett opptil verdiene fra de to australske studiene. Det tydeligste avviket gjelder Studentautonomi, hvor resultatet ligger klart under for de norske versjonene der denne skalaen er inkludert; for Klare mål ligger korrelasjonskoeffisientene noe under for EMS30 og EMS26.

EMS og læringstilnærminger/studiestrategier

Students læringstilnærminger og valg av studiestrategier danner et 'tradisjonelt' eksternt valideringskriterium for CEQ. Dokumentasjon av empiriske sammenhenger mellom students læringstilnærminger og deres opplevelse (perception) av undervisning og læringskontekst var en av hovedintensjonene i det andre Lancasterprosjektet (jf. del 1). Både CPQ, forløperen til CEQ, og Approaches to Studying Inventory (ASI) ble utviklet for dette formålet; denne ASI-versjonen danner også utgangspunkt for

senere versjoner av ASI (jf. Entwistle & McCune 2004; Pettersen 2004; Richardson 2000).

Man hadde, som tidligere nevnt, håpet å påvise sterkere empiriske sammenhenger enn det man faktisk oppnådde mellom studenters læring, basert på ASI, og opplevd undervisningskvalitet basert på CPQ. Det tydeligste resultatet var en betydelig *negativ* korrelasjon mellom opplevd arbeidsbyrde og *overflatetilnærming*, og en svakere *positiv* korrelasjon mellom *dybdetilnærming* og positive evalueringer av undervisning generelt. I senere studier, der man har anvendt CEQ og mer robuste ASI-versjoner, er det dokumentert noe tydeligere og sterkere empiriske koblinger mellom disse forholdene (jf. Lawless & Richardson 2002; Lizzio et al. 2002; Prosser & Trigwell 1999; Richardson 2000, 2005).

Tre ulike norske versjoner av ASI ble administrert sammen EMS i noen av kullene/klasseene som deltok i studie A og C, på følgende måte:

1. I studie A svarte fem ulike klasser/kull (n=217) på en 32-items versjon av ASI. Denne varianten er en rekonstruert kortversjon av den originale ASI, opprinnelig presentert av Richardson (1990, se også Lawless & Richardson 2002; Richardson 2000). De 32 itemene er likt fordelt på to skalaer betegnet *overflate-* og *dybdeorientering* i denne versjonen.¹⁰ Dette utvalget fra studie A er betegnet *utvalg 1* i tabell 14 og 15.
2. I studie A ble en kortversjon av en nyere utgave av ASI, kalt ASSIST, *Approaches and Study Skills Inventory for Students*, administrert til et utvalg bestående av tre klasser (n=86). En norsk versjon av den fullstendige versjonen av ASSIST med 52 item er validert av Diseth (2001; jf. Tait, Entwistle & McCune 1997). *Kortversjonen* har 36 item fordelt på tre hovedtilnærminger, *overflate-* og *dybdetilnærming* samt *strate-*

¹⁰ Betegnelseene 'orientation' og 'approach' anvendes in noen grad synonymt i ulike ASI-versjoner (jf. Entwistle & McCune 2004). ASI32 er oversatt til norsk ved forfatteren, og data fra en valideringsstudie av denne versjon inngår i artikkelen "In Quest of Alternative 'Quick-and-Easy' Versions of the ASI: The Approaches to Studying Inventory in a Norwegian Context" (under arbeid, 2006), jf. Pettersen 2004.

gisk tilnærming; den omtales her som ASSIST36.¹¹ Det aktuelle utvalget betegnes *utvalg 2* i tabell 14, 15 og 16.

3. I studie C ble samme kortversjon, ASSIST36, administrert til det ene utvalget i denne studien (n=436) sammen med EMS26. Dette utvalget betegnes *utvalg 3* i tabell 14, 15 og 16.
4. I studie C ble en nykonstruert norsk 18-items versjon av ASI administrert til det andre utvalget i studien sammen med EMS26, n=440.¹² De 18 itemene er likt fordelt på hovedtilnærmingene *overflate-* og *dybde-*tilnærming samt *strategisk tilnærming*; denne versjonen betegnes ASI-18(P), og utvalget er betegnet *utvalg 4* i tabellene 14–16.

I tabell 14 og 15 er resultater fra Wilson et al.s (1997) studie valgt som referanse. Her ble en kortversjon av ASI med 12 item spesielt utviklet for denne studien anvendt for å vurdere CEQs kriterievaliditet; 6 item for henholdsvis dybde- og overflatetilnærming. I studien er alle koeffisienter større enn 0,15 signifikante for $p < 0.001$. Tabell 14 gir en oversikt over korrelasjonskoeffisienter (Parsons 'r') for EMS-/CEQ-skalaer og *dybdetilnærming*. Tabell 15 viser resultater for *overflatetilnærming*.

Tabell 14

EMS-/CEQ-skalaer	Korrelasjon: Dybdetilnærming – EMS-/CEQ-skalaer				
	Dybdetilnærming (deep approach)				
	Studie A Utvalg 1 EMS30/ ASI32	Studie A Utvalg 2 EMS30/ ASSIST36	Studie C Utvalg 3 EMS26/ ASSIST36	Studie C Utvalg 4 EMS26/ ASI18(P)	Wilson et al. CEQ36/ ASI
God undervisning	0,24**	0,39**	0,20**	0,18**	0,24
Klare mål	0,17**	0,25*	0,18**	0,03	0,12
Arbeidsbyrde	0,13*	0,07	0,08	0,07	0,07
Vurderingsformer	0,11*	0,25*	0,12*	0,12*	0,21
Studentautonomi	0,13*	0,00	0,11	0,16**	0,19
Generelle ferdigheter	--	0,36**	0,36**	0,30**	0,37
	N=217	N=86	N=436	N=440	N=2130

**signifikant for $p < 0,01$ / *signifikant for $p < 0,05$

¹¹ Både ASSIST og ASSIST36 er oversatt til norsk av forfatteren uavhengig av Diseths (2001) oversatte versjon, og er validert med et større utvalg høgstudenter. Begge inventoriene viser tilfredsstillende psykometrisk egenskaper i samsvar med Diseths (2001) og Entwistle, Tait & McCunes (2000) resultater.

¹² Jf. Pettersen 2004, 2006.

Korrelasjonsmønsteret for *tre* skalaer i EMS-utvalgene, GU, ARB og GF, samsvarer i hovedsak med resultatene i studien til Wilson et al. For *Klare mål* er korrelasjonen noe *høyere* i utvalg 1–3, men er jevnt over *lavere* i alle utvalg for Studentautonomi og Vurderingsformer, med utvalg 2 som unntak for VF ($r=0,25$).

Tabell 15

EMS-/CEQ-skalaer	Korrelasjon: Overflatetilnærming – EMS-/CEQ-skalaer				
	Overflatetilnærming (surface approach)				
	Studie A Utvalg 1 EMS30/ ASI32	Studie A Utvalg 2 EMS30/ ASSIST36	Studie C Utvalg 3 EMS26/ ASSIST36	Studie C Utvalg 4 EMS26/ ASI18(P)	Wilson et al. CEQ36/ ASI
God undervisning	-0,20**	-0,37**	-0,16**	-0,12*	-0,34
Klare mål	-0,21**	-0,13	-0,26**	-0,23**	-0,29
Arbeidsbyrde	-0,46**	-0,25*	-0,48**	-0,51**	-0,48
Vurderingsformer	-0,27*	-0,43*	-0,37**	-0,26**	-0,47
Studentautonomi	-0,22*	-0,18	-0,16**	-0,22**	-0,29
Generelle ferdigheter	–	–	-0,14**	-0,04	-0,20
	N=217	N=86	N=436	N=440	N=2130

**¹) signifikant for $p<0,01$ / *¹) signifikant for $p<0,05$

I tabell 15 er det best samsvar mellom EMS-studiene og Wilson et al.s studie for korrelasjon mellom overflatetilnærming og *Arbeidsbyrde*, der- nest for *Vurderingsformer*. De øvrige korrelasjonskoeffisientene er stort sett noe lavere enn i Wilson et al.s studie.

Imidlertid, når tabell 14 og 15 vurderes samlet, støtter resultatene fra EMS-studiene i hovedsak Wilson et al.s konklusjoner, hvor de peker på at selv om “a number of the correlations were small, ... *the result are nevertheless consistent with a pattern of previous findings* at both an individual and course level in secondary and tertiary contexts using the CEQ, ... Clearly, then, the CEQ is measuring *aspects of the learning environment which are systematically associated with students’ reported learning processes*” (1997:43; min kursivering).

Et typisk trekk i dette totalmønsteret er *positiv* korrelasjon mellom *dybdetilnærming* og *Generelle ferdigheter* ($r\approx 0,35$) og *God undervisning* ($r\approx 0,20-0,25$). Korrelasjonen med *Arbeidsbyrde* er i praksis null; for de øvrige tre skalaene ligger koeffisientene stor sett i området $0,10-0,25$.

På den andre siden korrelerer overflatetilnærming *negativt* med samtlige skalaer i EMS/CEQ: Det er klarest sammenheng mellom overflatetilnærming og *Arbeidsbyrde*, $r \approx \pm 0,40-0,50$; dernest med *Vurderingsformer*, $r \approx \pm 0,30-0,40$. For *Studentautonomi* og *God undervisning* ligger korrelasjonen jevnt over i området $r \approx \pm 0,15-0,35$.

Både ASSIST36 og ASI18(P) omfatter skalaen/faktoren *strategisk tilnærming*. Den er definert som *prestasjonsmotivasjon* (intensjoner om ”å gjøre det godt”) og evne til *tidsplanlegging og organisering av eget studiearbeid*; forhold som gjelder læringens motivasjonelle sider – også omtalt som ”the will to learn”. Tabell 16 viser resultater for strategisk tilnærming som kriterium i utvalg 2, 3 og 4.

Tabell 16

EMS-/CEQ-skalaer	Korrelasjon: Strategisk tilnærming – EMS-/CEQ-skalaer		
	Strategisk tilnærming (strategic approach)		
	Studie A Utvalg 2	Studie C Utvalg 3	Studie C Utvalg 4
	EMS30/ASSIST36	EMS26/ASSIST36	EMS26/ASI18(P)
God undervisning	0,32**	0,22**	0,15**
Klare mål	0,38**	0,27**	0,12*
Arbeidsbyrde	0,12	0,01	0,09
Vurderingsformer	0,32**	0,07	0,04
Studentautonomi	0,10	0,03	0,02
Generelle ferdigheter	—	0,27**	0,26**
	N=86	N=436	N=440

**) signifikant for $p < 0,01$ / *) signifikant for $p < 0,05$

Korrelasjonen mellom strategisk tilnærming og *God undervisning* og *Klare mål* vises tydeligst i utvalgene der ASSIST36 er anvendt. I utvalg 2 er korrelasjonen mellom GU og KM hhv. 0,32 og 0,38; tilsvarende for utvalg 3: $r=0,22$ og $0,27$. I utvalg 4, med ASI18(P), er sammenhengene svakere, hhv. $r=0,15$ og $r=0,12$.

Korrelasjonen med *Generelle ferdigheter* for utvalg 3 og 4 er hhv. 0,27 og 0,28. De øvrige skalaene i EMS26 er i praksis ukorrelerte med strategisk tilnærming ($r \approx 0,0$); med unntak av *Vurderingsformer* i utvalg 2, $r=0,38$. Det kan tenkes at denne sammenhengen reflekterer et kontekstspesifikt forhold, særegent for utdanningen som utgjør utvalg 2.

Læringstilnærmingene 'strategisk', 'overflate' og 'dybde' bør sees i sammenheng; flere studier viser at de sammen danner ulike, typiske læringsmønstre som kan kobles til hvordan studenter innretter seg og presterer i akademiske kontekster (Entwistle 2000; Pettersen 2004). Et typisk mønster – *deep, strategic approaches to studying* – kjennetegnes ved en relativ klar dybdetilnærming (høy skåre) i forhold til overflatetilnærming (lav skåre), sammen med en relativt høy skåre på strategisk tilnærming. Mønstret forbindes generelt med akademisk orienterte studenter, også beskrevet som *selvregulerte og autonome* studenter. De gjør det vanligvis godt og lykkes i sine studier. Ut fra de samlede korrelasjonsmønstrene i tabell 14, 15 og 16, er det rimelig å anta at denne studentgruppen også skårer høyere på EMS enn studenter med motsatt læringsmønster, omtalt som *surface, apathetic approaches to studying*, eller *minimalistisk læringsorientering*. Dette mønstret kjennetegnes ved sterk preferanse for overflatetilnærming (høy skåre) i forhold til dybdetilnærming (lav skåre), sammen med en forholdsvis lav skåre for strategisk orientering. Denne studentgruppen møter oftere problemer i studiene, og de lykkes i mindre grad fullt ut i studiesituasjonen. Et typisk trekk er at gruppen skårer signifikant lavere på skalaene Arbeidsbyrde, Vurderingsformer – og i noen grad på Studentautonomi og God undervisning (jf. Entwistle 2000).

Generelle ferdigheter som eksternt kriterium

Skalaen *Generelle ferdigheter* bryter som nevnt med CEQs opprinnelige teoretiske faktorstruktur (s. 58f). Skalaen uttrykker studentenes vurdering av hvordan studiet bidrar til tilegnelse av bestemte *læringsresultater*; i dette tilfellet fagovergripende ferdigheter og generelle kompetanser, mens de øvrige CEQ-skalaene gjelder *opplevd kvalitet* i forhold til undervisnings- og læringssituasjonens *prosessuelle* aspekter (jf. Sadlo & Richardson 2003). Dermed framstår Generelle ferdigheter som *eksternt kriterium* (læringsprodukt) i forhold til opplevd undervisningskvalitet, slik dette

kommer til uttrykk på de øvrige CEQ– skalaene (jf. Lizzio et al. 2002; Wilson et al. 1997).

Tabell 17 viser korrelasjon mellom Generelle ferdigheter og øvrige skalaer i studie B og C, som omfatter EMS–versjoner med GF–skalaen. Resultater fra Wilson et al. (1997) er referert som sammenlikning; høyeste og laveste koeffisient fra de tre undersøkelsene som inngår i studien er gjengitt, hvor alle koeffisienter større enn 0,15 er signifikante for $p < 0.001$.

Tabell 17 Korrelasjon: *Generelle ferdigheter* – øvrige EMS–/CEQ–skalaer

	Studie B EMS30–GF	Studie C EMS26	Wilson et al.(1997) CEQ23 og CEQ36
EMS–/CEQ–skalaer	N=317	N=876	N≈10.600
God undervisning	0,42**	0,41**	0,40–0,46
Klare mål	0,33**	0,23**	0,29–0,33
Arbeidsbyrde	0,01	0,07	0,02–0,16
Vurderingsformer	0,23**	0,27**	0,20–0,37
Studentautonomi	--	0,28**	0,40–0,41

**) signifikant for $p < 0.01$

Korrelasjonskoeffisientene for studie B og C i tabell 17 ligger på samme nivå som resultatene i de australske studiene; den tydeligste forskjellen er en noe *svakere* sammenheng mellom *Generelle ferdigheter* og *Studentautonomi* for EMS26, $r=0,28$; her er laveste koeffisient hos Wilson et al. 0,40.

Valideringsstudiene A, B og C oppsummert – konklusjon

De tre oversatte versjonene av Course Experience Questionnaire – EMS30, EMS30–GF og EMS26 – viser gjennomgående tilfredsstillende psykometriske egenskaper. Resultatene er stort sett i tråd med resultater fra flere valideringsstudier utført i angloamerikanske utdanningskontekster, men med noe lavere reliabilitet for de kortere skalaene i EMS26. Det gjelder særlig Vurderingsformer og Studentautonomi, som også har lavere alfa-verdier i EMS30 og EMS30–GF. Likevel, EMS vurderes som et rimelig robust evalueringsverktøy vel egnet til å kartlegge *norske* studenters opplevelse og oppfatning av undervisningskvalitet i *norske utdanningskontekster*.

De mer detaljerte analysene av EMS' *begrepsvaliditet* (faktorstruktur) viser at norske høgstskolestudenter gjennomgående responderer på enkelt-utsagn i tråd med britiske og australske studenters responsmønster. Det kommer også til uttrykk på utsagn som avviker fra EMS/CEQs teoretiske, latente faktorstruktur. I EMS-studiene tilsvarer mønstret for disse jevnt over resultatene i tilsvarende studier med CEQ, og underbygger CEQs kryssvaliditet. Med andre ord fungerer EMS i all hovedsak på samme måte i norske høgstskolekontekster som CEQ anvendt ved angloamerikanske universiteter.

Analysene av EMS' *kriterievaliditet* underbygger CEQs kryssvaliditet eller "portability" ytterligere; også her er resultatene gjennomgående i samsvar med internasjonale studier. Samtlige studier viser at skalaene i EMS/CEQ er systematisk knyttet til relevante forhold som gjelder *undervisning og læring*: (a) studentenes fornøydhet med undervisning og læringsmiljø; (b) studentenes sjørrapportering når det gjelder læringstilnærminger og studiemønstre; (c) studentenes vurdering av undervisningens relevans for utvikling av generelle ferdigheter og fagovergripende kompetanser. Disse koblingene bygger på korrelasjonsanalyser som primært sier noe om styrken på sammenhengene, og mindre om kausale forhold, dvs. hva som påvirker hva, og hvilken retning relasjonene har.

Lizzio et al.s (2002) analyse av data fra CEQ36 og ASI basert på sti- og regresjonsanalyser gir et mer detaljert bilde av de kausale forholdene:

Hvordan faktorene påvirker hverandre både direkte og indirekte. Deres hovedfunn bekrefter og utdyper hovedmønstrer som tegnes i flere studier med CEQ:

1. Når studenter opplever *høy* arbeidsbyrde og *lite* relevante/egne vurderingsformer, påvirker dette dem i retning av *overflatetilnærminger* til læring og studiearbeid. Opplevd arbeidsbyrde har ikke en tilsvarende systematisk effekt på utviklingen av dybdetilnærminger. Imidlertid har opplevelsen av høy arbeidsbyrde en *indirekte negativ effekt* (mediert via overflatetilnærming) på utvikling av generelle ferdigheter.
2. Når studenter totalt sett opplever undervisnings- og læringsmiljøet som godt, påvirkes de til å ta i bruk dybdeorienterte lærings- og studiestrategier – og motsatt: Opplever læringsmiljøet som dårlig, motiverer det studenter til å anvende overflatetilnærminger. Det er opplevd undervisningskvalitet (GU) og opplevelsen av relevante og godt tilpassede vurderingsformer (VF) som gir sterkest predikasjon i forhold til valg og utvikling av *dybdetilnærminger* i studiesituasjonen.
3. Basert på data som gjelder studentenes tidligere skoleprestasjoner og eksamensresultater fra universitetsstudier, viser Lizzio et al.s studie følgende hovedtrekk:
 - Studenters opplevelse av kvaliteten på undervisnings- og læringsmiljøet påvirker resultater og prestasjoner både *direkte* og *indirekte*. Den *direkte effekten* går rett og slett ut på at studenter som opplever undervisnings- og læringsmiljøet som godt, presterer godt – og oppnår bedre resultater enn studenter som opplever sitt læringsmiljø som dårlig (jf. Marsh 1987). Den *indirekte effekten* går primært via utviklingen av dybdetilnærming, som ytterligere påvirker akademiske prestasjoner i positiv retning.
 - Studentenes opplevelse av kvaliteten på undervisnings- og læringsmiljøet gir *bedre* predikasjon med hensyn til framtidige akademiske prestasjoner enn tidligere skoleprestasjoner og eksamensresultater.

Lizzio et al. (2002:44) konkluderer med at studentevaluering av undervisnings- og læringsmiljø med CEQ gir grunnlag for didaktisk optimisme i høyere utdanning; resultatene fra et bredt utvalg studier bekrefter at “ ... *elements of the learning environment which are under teacher control can, and do, positively influence both the way students approach their study, and the learning outcomes they may achieve. Thus, interventions, if appropriately conceived, can and will ‘make a difference’* “ (min kursivering).

Oppsummering – kort sluttord

Course Experience Questionnaire i norsk oversettelse, *Erfaringer med studiet* (EMS), dokumenterer gjennomgående akseptable psykometriske egenskaper. Generelt er det en fordel å anvende instrumenter med adekvat validitet og reliabilitet; det styrker tiltroen til de resultater man oppnår. I evalueringssammenheng er imidlertid psykometrisk kvalitet ofte lite påaktet; i mange situasjoner er "... conclusions regarding the quality ... based on questionnaires about which even elementary measurement information is lacking" (Schmidt et al. 1995).

De ulike EMS-versjonene har kjente og rimelig robuste psykometriske kvaliteter, og egner seg slik sett som standardiserte evalueringsverktøy i norske utdanningskontekster. Svakheter som hefter ved noen av skalaene i de ulike versjonene, er dokumentert i de tre valideringsstudiene, A, B og C. Når hovedvekten legges på psykometriske forhold, er det EMS-26 som bør være et førstevalg. Flere valideringsstudier av denne versjonen i andre utdanningskontekster vil kunne vise hvor holdbar denne anbefalingen er generelt sett.

EMS kartlegger dimensjoner ved opplevd undervisningskvalitet og læringsmiljø som prinsipielt ligger under lærerkontroll, eller innenfor lærergruppers handlingsfelt. Dermed kan EMS fungere som supplement til læreres egne utviklingsorienterte evalueringsstrategier, som en komplementær strategi: EMS kan gi lærere meldinger og informasjon om ting de "trenger å høre", og som det kan være verdt å undersøke nærmere (jf. Patrick 1999). Dette poenget knytter grunnleggende sett an til læreres forestillinger og tenkning om hvordan undervisningen bør være, og hva som er standarder for god undervisning, trass i poengets paternalistiske undertone. Læreres oppfatninger og tenkning om undervisning er et sentralt element i Handal og Lauvås' (1983) begrep *praksisteori* (jf. Handal 1996; se også: Hativa & Goodyear 2002; Lauvås & Handal 2000; Wittek 2006).

Teoretisk spiller den enkelte lærers praksisteori en avgjørende rolle for hvordan læreren planlegger og gjennomfører sin undervisning. Og i forhold til studentevaluering danner praksisteorien basis for så vel tolkning

av evalueringsresultater som utforming av de evalueringsstrategier læreren velger å anvende. Slik sett kan studentevaluering i prinsippet tendere mot at ”som man roper i skogen, får man svar”. Derfor: en viktig betingelse for at studentevaluering skal fungere som feedback, som bidrar til *reell utvikling* av undervisning og læringsmiljø, er at lærerne blir ”seg bevisst hvordan en selv tenker «bak» den utformingen en gir undervisningen sin ... Da kan en se evalueringen – og drøftninger med andre om den – som en mulighet for å få bekreftelser på eller *utfordringer til sin egen praksisteori*” (Handal 1996:29ff; min kursivering).

Det er et særlig anliggende i denne rapporten å argumentere for EMS som et utviklingsorientert verktøy; EMS-resultater kan representere slike utfordringer til den enkelte lærers praksisteori. EMS-data kan bidra til at både den enkelte lærer og lærergrupper i større grad fokuserer på og reflekterer nærmere over dimensjoner ved undervisning og læringsmiljø som nødvendigvis ikke ligger fremst i bevisstheten, verken i den enkeltes eller gruppas bevissthet, eller den enkeltes praksisteori og lærerstabens ”kollektive” praksisteori, om man vil.

Course Experience Questionnaire er primært utviklet til bruk på kurs- og studieprogramnivå; evalueringsresultater refererer derfor til den involverte *lærergruppas*, og indirekte til *den enkelte lærers*, ”teaching effectiveness”. På den annen side handler CEQ-/EMS-resultater om ”course quality”, om hvordan kurs og programmer er organisert og gjennomføres. Det kan være klokt å ha denne distinksjonen klart i mente; for eksempel kan det tenkes at lærere *hver for seg* ikke oppfatter det som at de pålegger studentene en unødig stor og urimelig arbeidsbyrde i sine kurs, eller den delen av programmet de deltar i. Likevel kan det tenkes at studentene opplever det motsatt, og at dette viser seg i en påfallende lav skåre på skalaen Arbeidsbyrde på programnivå. Og videre kan det tenkes at lærerne, etter å ha undersøkt dette forholdet nærmere i fellesskap, også finner at summen av studie- og arbeidsoppgaver faktisk gir studentene en reell grunn til å oppleve at arbeidsbyrden er i største laget.

Selv om CEQ/EMS i hovedsak handler om forhold som faller innenfor lærernes umiddelbare handlingsrom, vil eksamens- og vurderingsformer,

som skalaen Vurderingsformer viser til, i mange sammenhenger framstår mer som en rammefaktor; som forhold den enkelte lærer må tilpasse undervisningen til. Eksamen er som regel en faktor som ikke endres i ”tide og utide” på initiativ fra enkeltlærere.

I EMS er disse forholdene både snevert og relativt generelt definert på samme tid. Skalaen *Vurderingsformer* i EMS26 består av tre utsagn som handler om hvorvidt vurderingsformene etterspør faktakunnskaper, og om man kan greie seg i studiet bare med mekanisk pugg og memorering av lærestoffet. Perspektivet utvides noe ved å trekke inn de to utsagnene fra skalaen *God undervisning* som gjelder mengden og kvaliteten på lærernes tilbakemeldinger og *formative vurderinger* (utsagn 7 og 17; jf. del 2).

Flere forhold bør motivere til å undersøke nærmere hvordan EMS kan styrkes som evalueringsverktøy når det gjelder eksamens- og vurderingsformer. For det første er Vurderingsformer en av de svakere skalaene psykometrisk sett, i samtlige EMS-studier.

For det andre representerer så vel *summativ* som *formativ* vurdering sterke drivkrefter i studenters læringsarbeid, ikke minst i forhold til den tiden studenter bruker på studiearbeidet. Videre er læringseffekter av *formativ* vurdering et tema som er løftet fram og tematisert internasjonalt; i norsk sammenheng er forholdet særlig aktualisert gjennom Kvalitetsreformens krav om hyppigere tilbakemeldinger, tettere oppfølging og mer vekt på skriftlige innleveringsoppgaver, samt anbefalinger om å innføre alternative eksamensformer.

Én innfallsvinkel med tanke på å styrke og videreutvikle vurderingsdimensjonen i EMS finner vi i et større prosjekt i Storbritannia under ledelse av Graham Gibbs og kollegaer. Prosjektet sikter på å kartlegge kritiske trekk ved institusjoners *vurderingsmiljø* og *-kultur* i forhold til studenters læring. I tilknytning til prosjektet har man utviklet *The Assessment Experience Questionnaire* (AEQ) (jf. Brown, Gibbs & Glover 2003; Gibbs 1999; Gibbs & Simpson 2003, 2004–5; Gibbs, Simpson & Macdonald 2003; Gibbs et. al 2006).

Valideringsstudier av AEQ i norske utdanningskontekster vil vise om dette spørreskjemaet innholder utsagn (item) og eventuelle delskalaer som med fordel kan supplere EMS' nåværende skalastruktur. Dette representerer en av flere mulige strategier som bør forfølges, for å styrke skalaen *Vurderingsformer* i EMS, som sentral dimensjon i studenters samlede opplevelse og vurdering av undervisning og læringsmiljø.

Referanser

- Ainley J 2001. *Course Experience Questionnaire 2000*. Victoria: Graduate Careers Council of Australia Ltd.
- Barrows H S 1985. *How to Design a Problem-Based Curriculum for the Preclinical Years*. New York: Springer Publishing Company.
- Bligh J, G Lloyd-Jones & G Smith 2000. Early effects of a new problem-based clinically oriented curriculum on students' perceptions of teaching, *Medical Education*, 34:487–89.
- Bowden J & F Marton 2004. *The University of learning*. 2nd ed. London: Routledge Falmer.
- Broomfield D & J Bligh 1998. An evaluation of the 'short form' course experience questionnaire with medical students. *Medical Education*, 32:367–369.
- Brown E, G Gibbs & C Glover 2003. Evaluation tools for investigating the impact of assessment regimes on student learning. *BEE-j Volume 2*, November 2003.
- Brown G & M Atkins 1988. *Effective Teaching in Higher Education*. London: Routledge.
- Bryman A & D Cramer 2005. *Qualitative Data Analysis with SPSS 12 and 13*. London: Routledge.
- Byrne M & B Flood 2003. Assessing the Teaching Quality of Accounting Programmes: an Evaluation of the Course Experience Questionnaire, *Assessment & Evaluation in Higher Education*, 28 (2):135–145.
- Cattell R B 1966. The Scree Test for the Number of Factors. *Multivariate Behavioral Research*, 1:245–276.
- Christoffersen K–A 2004. *Databehandling og statistisk analyse med SPSS*. 2.utg. Oslo: Unipub.
- Curtis D D & J P Keeves 2000. The Course Experience Questionnaire as an Institutional Performance Indicator. *International Education Journal*, 1 (2); <http://iej.cjb.net>
- Diseth Å 2001. Validation of a Norwegian version of the Approaches and Study Skills Inventory for Students (ASSIST): application of structural equation modelling. *Scandinavian Journal of Educational Research*, 45 (4):381–94.
- Diseth Å, S Pallesen, A Hovland & S Larsen 2006. Course experience, approaches to learning and academic achievement. *Education + Training*, 48 (2/3):156–169.
- Entwistle N J 1987. A Model of the Teaching–Learning Process. I: J T E Richardson & M W Eysenck (eds): *Student Learning: research in education and cognitive psychology*. Milton Keynes: SRHE and Open University Press.
- Entwistle N J 1988. Motivational Factors in Students' Approaches to Learning. I: R R Schmeck (ed): *Learning Strategies and Learning Style*. New York: Plenum Press.
- Entwistle N J 1997. Contrasting perspectives on learning. I: F Marton, D Hounsell & N J Entwistle (eds): *The experience of learning. Implications for teaching and studying in higher education*, 2nd ed. Edinburgh: Scottish Academic Press.

- Entwistle N J 2000. Approaches to Studying and Levels of Understanding: The Influences of Teaching and Assessment. I: J C Smart (ed): *Higher Education: Handbook of Theory and Research, Volume XV*. New York: Agathon Press.
- Entwistle N J 2005. Ways of thinking and ways of teaching across contrasting subject areas. Paper to be presented at the ISL2005 conference on "Improving Student Learning by Assessment", London, 5.–7. September 2005, se: <http://www.ed.ac.uk/etl/publications.html#occasionalpapers>. Besøkt 3. september 2006.
- Entwistle N J & V McCune 2004. The Conceptual Base of Study Strategy Inventories. *Educational Psychology Review*, 16 (4):325–345.
- Entwistle N J & P Ramsden 1983. *Understanding Student Learning*. London: Croom Helm.
- Entwistle N J & H Tait 1990. Approaches to learning, evaluations of teaching, and preferences for contrasting academic environments. *Higher Education*, 19:169–194.
- Entwistle N J & J D Wilson 1977. *Degrees of Excellence: The Academic Achievement Game*. Kent: Hodder and Stroughton.
- Entwistle N J, V McCune & P Walker 2001. Conceptions, Styles, and Approaches Within Higher Education: Analytic Abstractions and Everyday Experience. I: R J Sternberg & L–F Zang (Eds): *Perspectives on Thinking, Learning, and Cognitive Styles*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Entwistle N J, H Tait & V McCune 2000. Patterns of response to an Approaches to Studying Inventory across contrasting groups and contexts. *European Journal of the Psychology of Education*, 15:33–48.
- Espeland V & O Indrehus 2003. Evaluation of students' satisfaction with nursing education in Norway, *Journal of Advanced Nursing*, 42 (3):226–236.
- Gallagher M 1999. The CEQ delivering higher quality outcomes for the higher education section. I: T Hand & K Trembart (eds) 1999. *The Course Questionnaire Symposium 1998*. Australia: The Department of Education, Training and Youth Affairs.
- Gibbs G 1999. Using assessment strategically to change way students learn. I: S Brown & A Glasner (Eds): *Assessment Matters in Higher Education*. Buckingham: SRHE and Open University Press.
- Gibbs G & C Simpson 2004–5. Conditions Under Which Assessment Supports Students' Learning. *Learning and Teaching in Higher Education*, Issue 1.
- Gibbs G, C Simpson & R Macdonald 2003. *Improving student learning through changing assessment – a conceptual framework*. EARLI Conference, Padova, Italy.
- Gibbs G, H Dunbar–Goddet, C Simpson & C Rust 2006. *Characterising features of programme level assessment environments that support learning*. Improving Student Symposium, University of Bath, England, September, 2006.
- Ginns P 2003. *Scale Structure of the Student Course Experience Questionnaire over Field and Level of Study*. HERDSA 2003 conference proceedings, jf. <http://surveys.canterbury.ac.nz/herdsa03/pdfsnon/N1066.pdf>

- Hair J F, R E Anderson, R L Tatham & W C Black 1998. *Multivariate data analysis*, 5th ed. Upper Saddle River: Prentice Hall.
- Hand T & K Trembart (eds) 1999. *The Course Questionnaire Symposium 1998*. Australia: The Department of Education, Training and Youth Affairs.
- Handal G 1996. *Studentevaluering av undervisning. Håndbok for lærere og studenter i høyere utdanning*. Oslo: Cappelen Akademisk Forlag.
- Handal G 2006. Studentevaluering av undervisning. I: H Strømsø, K Hofgaard Lycke & P Lauvås (red): *Når læring er det viktigste. Undervisning i høyere utdanning*. Oslo: Cappelen Akademisk Forlag.
- Handal G & P Lauvås 1983. *På egne vilkår. En strategi for veiledning med lærere*. Oslo: Cappelen.
- Hativa N & P Goodyear (eds) 2002. *Teacher Thinking, Beliefs and Knowledge in Higher Education*. London: Kluwer Academic Publishers.
- Howitt D & D Cramer 2005. *Introduction to SPSS in Psychology*, 3.rd ed. Essex: Pearson Education Ltd.
- Kim J-O & C W Mueller 1978. *Factor analysis. Statistical methods and practical issues*. Beverly Hills: Sage.
- Kreber C 2003. The Relationship between Students' Course Perception and their Approaches to Studying in Undergraduate Science Courses: A Canadian experience. *Higher Education Research & Development*, 22 (1):57–75.
- Lauvås P & G Handal 2000. *Veiledning og praktisk yrkesteori*. 2.utg. Oslo: Cappelen Akademisk Forlag.
- Lauvås P & A Jakobsen 2002. *Exit eksamen – eller? Former for summativ evaluering i høgre utdanning*. Oslo: Cappelen Akademisk Forlag.
- Lawless C J & J T E Richardson 2002. Approaches to studying and perceptions of academic quality in distance education. *Higher Education*, 44:257–282.
- Linke R D 1991. *Report of the research group on performance indicators in higher education*. Canberra: AGPS.
- Lizzio A, K Wilson & R Simons 2002. University students' perceptions of the learning environment and academic outcomes: Implications for theory and practice. *Studies in Higher Education*, 27 (1):27–52.
- Loelin J C 2004. *Latent Variable Models. An Introduction to Factor, Path, and Structural Equation Analysis*, 4th ed. Mahwah, NJ: Lawrence Erlbaum.
- Lyon P M & G D Hendry 2002. The Use of Course Experience Questionnaire as a Monitoring Evaluation Tool in a Problem-based Medical Programme. *Assessment & Evaluation in Higher Education*, 27 (4):339–52).
- Marsh H W 1987. Students' evaluations of university teaching: Research findings, methodological issues, and direction for further research. *International Journal of Education Research*, 11: Whole issue no 3.
- Marsh H W & L A Roche 1997. Making Students' Evaluations of Teaching Effectiveness Effective. *American Psychologist*, 52 (11):1187–1197.
- Marton F & R Säljö 1976. On qualitative differences in learning: I. Outcome and process. *British Journal of Educational Psychology*, 46:4–11.

- Marton F, D Hounsell & N Entwistle 1984. *The Experience of learning*. (2.utg 1997). Edinburgh: Scottish Academic Press.
- Marton F et al. 1977: *Inlärning och omvärldsuppfattning. En bok om den studerande människan*. Stockholm: AWE/Geber.
- McInnis C, P Griffin, R James & H Coates 2001. *Development of the Course Experience Questionnaire (CEQ)*. The University of Melbourne: Center for the Study of Higher Education and Assessment Research Center, Faculty of Education.
- Meyer J H F & P Parsons 1989. Approaches to studying and course perceptions using the Lancaster Inventory – a comparative study. *Studies in Higher Education*, 14: 137–153.
- NOU 1988:28. *Med viten og vilje*. Oslo: Kultur- og vitenskapsdepartementet.
- Nordvik H & P Ulleberg 2000. *Teststatistikk*. Trondheim: Tapir Akademisk Forlag.
- Pallant J 2005. *SPSS Survival Manual*, 2.nd ed. Maidenhead: Open University Press.
- Patrick, K 1999. Using CEQ results at RMIT. I: T Hand & K Trembart (eds) 1999. *The Course Questionnaire Symposium 1998*. Australia: The Department of Education, Training and Youth Affairs.
- Patrick K 2003. The CEQ in practice: using the CEQ for improvement. Paper presented by invitation at the GCCA symposium, Canberra, 24–25 March 2003.
- Pedhazur E J & L P Schmelkin 1991. *Measurement, design, and analysis. An integrated approach*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Pettersen R C 2004. Studenters lærings- og studiestrategier: Kvalitetsindikatorer i høyere utdanning? *Uniped*, 27 (2):44–65.
- Pettersen R C 2005. *Kvalitetslæring i høyere utdanning. Innføring i problem- og praksisbasert didaktikk*. Oslo: Universitetsforlaget.
- Pettersen R C 2006. In Quest of Alternative “Quick-and-Easy” Versions of the ASI: The Approaches to Studying Inventory in a Norwegian Context (under arbeid).
- Prosser M & K Trigwell 1999. *Understanding Learning and Teaching: The Experience in Higher Education*. Buckingham: SRHE and Open University Press.
- Ramsden P 1991. A performance indicator of teaching quality in higher education: the Course Experience Questionnaire. *Studies in Higher Education*, 20:129–150.
- Ramsden P 1992/1999. *Learning to teach in higher education* (2nd ed. 2003). London: Routledge (dansk utgave 1999: *Strategier for bedre undervisning*. København: Gyldendal).
- Ramsden P, E Martin & J Bowden 1989. School environment and sixth form pupils’ approaches to learning. *British Journal of Educational Psychology*, 59:129–142.
- Richardson J T E 1990. Reliability and replicability of the Approaches to Studying Questionnaire. *Studies in Higher Education*, 15:155–168.
- Richardson J T E 1994. A British evaluation of the Course Experience Questionnaire. *Studies in Higher Education*, 19 (1):59–68
- Richardson J T E 2000. *Researching Student Learning*. London: SRHE and Open University Press.

- Richardson J T E 2004. Methodological Issues in Questionnaire-Based Research on Student Learning in Higher Education. *Educational Psychology Review*, 16 (4): 347–358.
- Richardson J T E 2005. Students' perceptions of academic quality and approaches to studying in distance education. *British Educational Research Journal*, 31 (1):7–27.
- Richardson J T E & L Price 2003. Approaches to studying and perceptions of academic quality in electronically delivered courses. *British Journal of Educational Technology*, 34 (1): 45–56.
- Richardson J T E & A Woodley 2001. Perceptions of academic quality among students with hearing loss in distance education. *Journal of Educational Psychology*, 93: 563–570.
- Riding R & S Rayner 1998. *Cognitive Styles and Learning Strategies. Understanding Style Differences in Learning and Behaviour*. London: David Fulton Publ.
- Sadlo G 1997. Problem-based learning enhances the educational experiences of occupational therapy students. *Education for Health*, 10 (10):101–114.
- Sadlo G & J T E Richardson 2003. Approaches to Studying and Perceptions of the Academic Environment in Students Following Problem-Based and Subject-Based Curricula. *Higher Education Research & Development*, 22 (3):253–274.
- Samuelsen F 2006. Studentevaluering av undervisning, mer enn å måle tilfredshet. *Uniped*, 29 (2): 41–55.
- Schmeck R R (ed) 1988. *Learning Strategies and Learning Style*. New York: Plenum Press.
- Schmidt H G, D Dolmans, W H Gijsselaers & J E Des Marchais 1995. Theory-Guided Design of a Rating Scale for Course Evaluation in Problem-Based Curricula. *Teaching and Learning in Medicine*, 7 (2):82–91.
- Smith H W 1981. *Strategies of social research. The methodological imagination*. Englewood Cliffs, NJ: Prentice-Hall.
- Steele, G A, S A West & D T Simeon 2003. Using A Modified Course Experience Questionnaire (CEQ) to Evaluate the Innovative Teaching of Communication Skills. *Education for Health*, 16 (2):133–144.
- Sternberg R J & L-F Zhang (Eds) 2001. *Perspectives on Thinking, Learning, and Cognitive Styles*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Stortingsmelding nr. 40 (1990–1991). *Fra visjon til virke. Om høyere utdanning*. Oslo: Utdannings- og forskningsdepartementet.
- Stortingsmelding nr. 27 (2000–2001). *Gjør din plikt – Krev din rett. Kvalitetsreform av høyere utdanning*. Kirke- Utdannings- og Forskningsdepartementet.
- Tait H, N J Entwistle & V McCune 1997. ASSIST: a reconceptualisation of the Approaches to Studying Inventory. In C. Rust, (ed.), *Improving Student Learning: Improving Students as Learners*. Oxford: Oxford Brookes University, Oxford Centre for Staff and Learning Development.
- Trigwell K & M Prosser 1991. Improving the quality of student learning: the influence of learning context and student approaches to learning on learning outcome. *Higher Education*, 22:251–266.

- Ulleberg P & H Nordvik 2001. *Faktoranalyse. Innføring i faktorteori og faktoranalyse*. Trondheim: Tapir Akademisk Forlag.
- Watkins D 2001. Correlates of Approaches to Learning: A Cross–Cultural Meta Analysis. I: R J Sternberg & L–F Zhang (Eds): *Perspectives on Thinking, Learning, and Cognitive Styles*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Webb G 1997. Deconstruction deep and surface: Towards a critique of phenomenography. *Higher Education*, 33:195–212.
- Wilson K, A Lizzio & P Ramsden 1997. The Development, Validation and Application of the Course Experience Questionnaire. *Studies in Higher Education* 22(1): 33–53.
- Wittek L 2006. Om undervisning og læring. I: H Strømsø, K Hofgaard Lycke & P Lauvås (red): *Når læring er det viktigste. Undervisning i høyere utdanning*. Oslo: Cappelen Akademisk Forlag.

Vedlegg

Vedlegg 1: EMS30

Vedlegg 2: EMS30–GF

Vedlegg 3: EMS26

Vedlegg 4: Skalalengde, inter–item korrelasjon og Cronbachs alfa

Alle utsagn merket * i vedlegg 1, 2 og 3 gjelder negativt formulerte utsagn som invertertes før videre databearbeiding.

Vedlegg 1

EMS30

Skala	Nr	
KM	1	Det er alltid lett å forstå og finne ut hvilke krav som stilles til studiearbeidet mitt
SA*	2	I studiet er det <i>få</i> muligheter til å fordype seg i temaområder ut fra egne interesser
GU	3	Lærerne mine motiverer meg til å gjøre mitt beste i studiet
ARB*	4	Arbeidsbyrden på utdanningen er for stor
VF*	5	Lærene gir ofte uttrykk av det <i>ikke</i> er noe de kan lære av studentene
KM	6	Vanligvis har jeg en klar formening om hvordan jeg ligger an og hva som forventes av meg i studiet
GU	7	Lærerne bruker mye tid på å kommentere og gi tilbakemeldinger på oppgavene mine
VF*	8	For å klare seg godt på dette studiet, er alt du egentlig trenger en god hukommelse
SA	9	Vi blir oppmuntret til å utvikle faglige interesser så langt som mulig i dette studiet
ARB*	10	Undervisningsplanene for studiet forsøker å dekke altfor mange emner og temaer
SA	11	I studiet står jeg ganske fritt til å legge opp egen læring slik det passer best for meg
VF*	12	Jeg synes lærerne er mer interessert i å teste hva vi husker enn hva vi har forstått
KM	13	Det er ofte vanskelig å finne ut nøyaktig hva som forventes av meg på studiet
ARB	14	Som regel synes jeg vi får nok tid til å forstå de tingene vi skal lære oss
GU	15	Lærerne legger virkelig vekt på å forstå vanskeligheter vi møter i studiearbeidet
SA	16	Vi har ganske stor frihet i forhold til studiearbeid og oppgaver vi arbeider med
GU	17	Lærerne gir nyttige tilbakemeldinger om hvordan jeg ligger an i studiet
GU	18	Lærerne våre er virkelig gode til å forklare tingene for oss
KM*	19	Verken hovedmål eller delmål for studiet er formulert klart og tydelig nok
GU	20	Lærerstabten anstrenger seg og jobber hardt for å gjøre lærestoffet interessant
VF*	21	Det er altfor mange av lærerne som bare stiller oss spørsmål om faktakunnskaper
ARB*	22	Det stilles høye krav og det er stort press på meg som student i dette studiet
VF*	23	Tilbakemeldinger på studiearbeidene mine gis vanligvis som korte skriftlige kommentarer og/eller karakterer
SA	24	Vi diskuterer ofte med lærerne hvordan vi bør arbeide for lære på best mulig måte
GU*	25	Lærerne viser <i>ingen</i> virkelig interesse for studentenes ideer, forslag og meninger
VF*	26	Man kan komme seg gjennom dette studiet bare ved å arbeide hardt i tiden før eksamen
GU	27	På utdanningen forsøker lærerne virkelig å få oss til å yte vårt beste
SA*	28	Studiet gir liten frihet og få valg når det gjelder eksamens- og vurderingsformer
KM	29	Lærerne her gjør det klart helt fra begynnelsen av, hva de forventer av oss studenter
ARB*	30	Omfanget av lærestoff/pensum gjør at vi ikke kan lære og forstå alt grundig
SPM	31	I det store og hele er jeg godt fornøyd med kvaliteten på utdanningen (studiet)

Vedlegg 2

EMS30–GF

Skala	Nr	
KM	1	Det er alltid lett å forstå og finne ut hvilke krav som stilles til studiearbeidet mitt
GF	2	Studiet bidrar til at jeg utvikler mine ferdigheter i samarbeid og som gruppedeltaker
GU	3	Lærerne mine motiverer meg til å gjøre mitt beste i studiet
ARB*	4	Arbeidsbyrden på utdanningen er for stor
VF*	5	Lærerne gir ofte uttrykk av det <i>ikke</i> er noe de kan lære av studentene
KM	6	Vanligvis har jeg en klar formening om hvordan jeg ligger an og hva som forventes av meg i studiet
GU	7	Lærerne bruker mye tid på å kommentere og gi tilbakemeldinger på oppgavene mine
VF*	8	For å klare seg godt på dette studiet, er alt du egentlig trenger en god hukommelse
GF	9	Studiet bidrar til at jeg utvikler mine ferdigheter i å organisere eget arbeid
ARB*	10	Undervisningsplanene for studiet forsøker å dekke altfor mange emner og temaer
GF	11	Jeg har utviklet mine ferdigheter i å kommunisere og uttrykke meg skriftlig ved å delta på dette studiet
VF*	12	Jeg synes lærerne er mer interessert i å teste hva vi husker enn hva vi har forstått
KM	13	Det er ofte vanskelig å finne ut nøyaktig hva som forventes av meg på studiet
ARB	14	Som regel synes jeg vi får nok tid til å forstå de tingene vi skal lære oss
GU	15	Lærerne legger virkelig vekt på å forstå vanskeligheter vi møter i studiearbeidet
GF	16	Studieopplegget bidrar til at jeg utvikler mine ferdigheter i problemløsning
GU	17	Lærerne gir nyttige tilbakemeldinger om hvordan jeg ligger an i studiet
GU	18	Lærerne våre er virkelig gode til å forklare tingene for oss
KM*	19	Verken hovedmål eller delmål for studiet er formulert klart og tydelig nok
GU	20	Lærerstaben anstrenger seg og jobber hardt for å gjøre lærestoffet interessant
VF*	21	Det er altfor mange av lærerne som bare stiller oss spørsmål om faktakunnskaper
ARB*	22	Det stilles høye krav og det er stort press på meg som student i dette studiet
VF*	23	Tilbakemeldinger på studiearbeidene mine gis vanligvis som korte skriftlige kommentarer og/eller karakterer
GF	24	Jeg utvikler mine analytiske evner og ferdigheter i dette studiet
GU*	25	Lærerne viser <i>ingen</i> virkelig interesse for studentenes ideer, forslag og meninger
VF*	26	Man kan komme seg gjennom dette studiet bare ved å arbeide hardt i tiden før eksamen
SA	27	Studentene gis stor valgfrihet i forhold til oppgaver og studiearbeid som skal utføres
GF	28	Studiet bidrar til at jeg blir sikrere på å håndtere nye og ukjente problemstillinger
KM	29	Lærerne her gjør det klart helt fra begynnelsen av, hva de forventer av oss studenter
ARB*	30	Omfanget av lærestoff/pensum gjør at vi ikke kan lære og forstå alt grundig
SPM	31	I det store og hele er jeg godt fornøyd med kvaliteten på utdanningen (studiet)

EMS26

Skala	Nr	
KM	1	Det er alltid lett å forstå og finne ut hvilke krav som stilles til studiearbeidet mitt
GF	2	Studiet bidrar til at jeg utvikler mine ferdigheter i samarbeid og som gruppedeltaker
GU	3	Lærerne mine motiverer meg til å gjøre mitt beste i studiet
ARB*	4	Arbeidsbyrden på utdanningen er for stor
SA*	5	I studiet er det <i>få</i> muligheter til å fordype seg i temaområder ut fra egne interesser
KM	6	Vanligvis har jeg en klar formening om hvordan jeg ligger an og hva som forventes av meg i studiet
GU	7	Lærerne bruker mye tid på å kommentere og gi tilbakemeldinger på oppgavene mine
VF*	8	For å klare seg godt på dette studiet, er alt du egentlig trenger en god hukommelse
GF	9	Studiet bidrar til at jeg utvikler mine ferdigheter i å organisere eget arbeid
SA	10	I studiet står jeg ganske fritt til å legge opp egen læring slik det passer best for meg
GF	11	Jeg har utviklet mine ferdigheter i å kommunisere og uttrykke meg skriftlig ved å delta på dette studiet
VF*	12	Jeg synes lærerne er mer interessert i å teste hva vi husker enn hva vi har forstått
KM	13	Det er ofte vanskelig å finne ut nøyaktig hva som forventes av meg på studiet
ARB	14	Som regel synes jeg vi får nok tid til å forstå de tingene vi skal lære oss
GU	15	Lærerne legger virkelig vekt på å forstå vanskeligheter vi møter i studiearbeidet
GF	16	Studieopplegget bidrar til at jeg utvikler mine ferdigheter i problemløsning
GU	17	Lærerne gir nyttige tilbakemeldinger om hvordan jeg ligger an i studiet
GU	18	Lærerne våre er virkelig gode til å forklare tingene for oss
GU	19	Lærerstaben anstrenger seg og jobber hardt for å gjøre lærestoffet interessant
VF*	20	Det er altfor mange av lærerne som bare stiller oss spørsmål om faktakunnskaper
ARB*	21	Det stilles høye krav og det er stort press på meg som student i dette studiet
GF	22	Jeg utvikler mine analytiske evner og ferdigheter i dette studiet
SA	23	Vi har ganske stor frihet i forhold til studiearbeid og oppgaver vi arbeider med
GF	24	Studiet bidrar til at jeg blir sikrere på å håndtere nye og ukjente problemstillinger
KM	25	Lærerne her gjør det klart helt fra begynnelsen av, hva de forventer av oss studenter
ARB*	26	Omfanget av lærestoff/pensum gjør at vi ikke kan lære og forstå alt grundig
SPM	27	I det store og hele er jeg godt fornøyd med kvaliteten på utdanningen (studiet)

Skalalengde, inter-item korrelasjon og Cronbachs alfa

Gjennomsnitt, Inter-item korrelasjon, r =	3 item	4 item	5 item	6 item	8 item	12 item	Vurdering av indre konsistens, Cronbachs α
0,450	0,71	0,77	0,80	0,83	0,87	0,91	Meget
0,400	0,67	0,72	0,77	0,80	0,84	0,89	høy –
0,350	0,62	0,68	0,73	0,76	0,81	0,87	høy
0,325	0,59	0,66	0,71	0,74	0,79	0,85	til
0,300	0,56	0,63	0,68	0,72	0,77	0,84	moderat –
0,250	0,50	0,57	0,63	0,67	0,73	0,80	til
0,225	0,47	0,54	0,59	0,63	0,70	0,78	lav –
0,200	0,43	0,50	0,56	0,60	0,67	0,79	mot grense
0,185	0,41	0,48	0,53	0,58	0,64	0,73	for akseptabel
0,170	0,38	0,45	0,51	0,55	0,62	0,71	reliabilitet
0,150	0,35	0,41	0,47	0,51	0,59	0,68	?
0,125	0,30	0,36	0,42	0,43	0,53	0,63	??
0,100	0,25	0,31	0,36	0,40	0,47	0,57	